ifo WORKING PAPERS

387
2022
December 2022

Redesigning Automated Market Power Mitigation in Electricity Markets

Jacqueline Adelowo, Moritz Bohland

ifo INSTITUTE
Leibniz Institute for Economic Research
at the University of Munich

Imprint:

ifo Working Papers
Publisher and distributor: ifo Institute – Leibniz Institute for Economic Research at the University of Munich
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49(0)89 9224 0, Telefax +49(0)89 985369, email ifo@ifo.de
www.ifo.de

An electronic version of the paper may be downloaded from the ifo website:
www.ifo.de

ifo Working Paper No. 387

Redesigning Automated Market Power Mitigation in Electricity Markets*

Abstract

Electricity markets are prone to the abuse of market power. Several US markets employ algorithms to monitor and mitigate market power abuse in real time. The performance of automated mitigation procedures is contingent on precise estimates of firms' marginal production costs. Currently, marginal cost are inferred from the past offers of a plant. We present new estimation approaches and compare them to the currently applied benchmark method. We test the performance of all the approaches on auction data from the Iberian power market. The results show that our novel approaches outperform the benchmark approach significantly, reducing the mean absolute estimation error from 11.53 €/MWh to 2.77 €/MWh for our most precise alternative approach. Applying this result to a market mitigation simulation we find sizeable overall welfare gains and welfare transfers from supplier to buyer surplus. Our research contributes to accurate monitoring of market power and improved automated mitigation. Although we focus on power markets, our findings are applicable to monitoring of renewable energy tenders or market power surveillance in rail and air traffic.

JEL code: D22, D43, D44, D47, L13, L94
Keywords: Regulation, automated mitigation procedure, best-response pricing, market power, Electricity, mark-up

Jacqueline Adelowo**
ifo Institute – Leibniz Institute for Economic Research at the University of Munich
Poschingerstr. 5
81679 Munich, Germany
adelowo@ifo.de

Moritz Bohland
Technical University Munich
Center for Energy Markets
Arcisstr. 21
80333 Munich, Germany

*We are grateful for helpful comments from our departments, especially Sebastian Schwenen, Hanna Hottenrott, and Valeriya Azarova, participants of the 10th Mannheim Conference on Energy and the Environment, and participants of the 49th EARIE Annual Conference.
** Corresponding author.

Declaration of interests The authors declare that they have no known competing conflict of interest or financial interests that could have appeared to influence the work reported in this paper.

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CRediT author statement Jacqueline Adelowo: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Visualization; Roles/Writing - Original Draft; Moritz Bohland: Conceptualization; Formal analysis; Investigation; Methodology; Visualization; Roles/Writing – Original Draft.

1. Introduction

The liberalization of power markets entailed efficiency gains and cost reductions for electricity producers (e.g. Newbery and Pollitt, 1997, Davis and Wolfram, 2012), but these gains did not necessarily translate into lower market prices (Newbery, 1997). The missing link between cost reductions for producers and reductions in power prices is, at least partially, attributed to market power abuse by electricity generating companies. Market power exertion in liberalized electricity markets is documented for a wide range of markets and periods (e.g. Green and Newbery, 1992, Borenstein et al., 1999, Ciarreta and Espinosa, 2010). Limited storage capacities, inelastic short-run demand, and high market concentration render power markets especially prone to market power exertion. As market power abuse is both inefficient and undesired by policy makers, regulators aim at mitigating undue market power.

Existing mitigation strategies include the implementation of price caps (Wilson, 2000), stringent application of antitrust policies (Green, 1996, Borenstein et al., 1999), fostering of vertical integration (Mansur, 2007, Bushnell et al., 2008), and the implementation of forward contracting obligations for suppliers (Allaz and Vila, 1993, de Frutos and Fabra, 2012). In several US markets, system operators go one step further and monitor and mitigate market power in real time. To that end, system operators implemented automated mitigation procedures (AMP), i.e. algorithms to screen all supply offers, detect undue market power, and mitigate affected offers. Future electricity systems will depend even more on flexible, quickly dispatchable generators at the margin to balance increasing shares of intermittent renewables (in absence of sufficient storage and short-term demand response) – hence, raising the risk of market power abuse. Graf et al. (2021) point out how this will heighten relevance of AMPs to work properly in increasingly decarbonized systems. A striking example of this is the recent power crisis with high marginal prices from natural gas-fired generation due to the Russo-Ukrainian war. This allows powerful firms with a diverse generation portfolio to strategically deploy their units to maximize windfall profits.

In this paper we contribute to improved algorithms for automated mitigation of market power in multi-unit uniform price auctions. In electricity markets, market power is typically measured by the difference between observed offers and underlying marginal cost of power production.[1] Therefore, marginal cost estimates should be as accurate as possible to ensure unbiased measurement of market power (Bushnell et al., 2008) and welfare-improving mitigation thereof. When all cost components of power production are known, engineering based bottom-up calculations deliver precise estimates of marginal cost. However, cost components and power plant characteristics are private information and firms have an incentive to overstate costs. Instead, system operators thus infer marginal cost of power plants from past offers of the respective plant, which leaves room for strategic manipulation by firms (Shawhan et al., 2011). We use this best-practise approach as a benchmark for further analysis and present alternative methods that deliver more accurate marginal cost estimates.

To test the accuracy of the benchmark approach and alternative methods, we employ micro-level bidding data from the Iberian day-ahead electricity market. First, we calculate marginal cost of power production bottom-up to obtain a measure for "true" marginal cost. To that end, we employ detailed information on power plant characteristics and all relevant cost components. In a second step, we test the benchmark approach based on past offers and compare the outcomes to the true marginal cost we derived in the first step. We then proceed by testing the accuracy of alternative estimation methods and assess their performance as compared to the benchmark approach currently employed by system operators. Using our preferred method, we carry out a mitigation simulation and welfare analysis on the data.

First, we test a theory-driven approach, which is based on Wolak (2003a, 2007) and accounts for the price reducing effect of a firm's forward obligations. We assume power pro-

[1] Going back to the Lerner-Index of the degree of monopoly power as $\frac{price - marginal\ cost}{price}$.

ducing companies to submit profit-maximizing offer curves as a best-response to the offers of competing firms. Under this assumption, we infer marginal cost of power production that justify observed offers. We designate this approach as "best-response" approach. Additionally, we present two approaches, which methodologically build on the benchmark approach used by system operators but address major flaws of the existing method. In the first of these two approaches, we additionally control for distortions caused by potential start-up and ramping cost. We refer to this approach as the "start-up" approach. The last estimation method we propose represents an extension to the start-up approach, where we now define clusters of similar power plants and estimate marginal cost for the whole cluster of plants. We refer to this method as the "clustering" approach.

The results of our empirical analysis reveal a low estimation accuracy of the currently applied benchmark approach. For the sample of power plants that we analyze, we find a mean absolute deviation of 11.53 €/MWh between marginal cost estimates following the benchmark approach and true marginal cost. All suggested alternative approaches deliver more precise estimates. Mean absolute deviations accrue to 8.92 €/MWh for the best-response approach, 7.27 €/MWh for the start-up approach, and merely 2.77 €/MWh for the clustering approach. The clustering approach does not only deliver the most precise estimates, but likewise limits the scope for strategic manipulation of estimates by firms. This is because estimates are based on past bids of a group of plants instead of just one plant. Strategic manipulation of estimates and thus mitigation would hence require a significant extent of coordination among firms. The therefore assess the risk of strategic manipulation as reduced. Applying the clustering approach to an AMP simulation on the data, we find sizeable overall welfare gains and welfare transfers from supplier to buyer surplus.

Our findings provide system operators with improved estimation techniques of power plants' marginal cost and with more accurate methods for monitoring and real time mitigation of market power. Equipped with precise marginal cost estimates, system operators can apply automated mitigation more stringently, and achieve increased market efficiency and reduced costs for consumers. At the same time, improved accuracy benefits producers as the scope for unjust mitigation of offers based on flawed marginal cost estimates is reduced. The main use cases for our approaches are automated procedures for market power mitigation in spot, balancing, and reserve electricity markets. Yet, the approaches can likewise find application in other markets, e.g. for monitoring in renewable energy tenders or price and market power surveillance in rail and air traffic. Additionally, marginal cost estimation approaches which are not contingent on private information facilitate power market research for scholars. The suggested approaches are especially valuable when a bottom-up calculation is infeasible due to limited accessibility of private information on cost components.

Considering the widespread application of AMPs in US power markets and the immediate effect of mitigation procedures on market prices, producer and consumer rents, as well as investment decisions, literature on AMPs is rather scarce and to a large extent of qualitative nature. Twomey et al. (2006) and García and Reitzes (2007) address AMPs in their reviews of market power monitoring and mitigation measures. Helman (2006) and Graf et al. (2021) assess and compare market power monitoring and mitigation procedures in several US markets. Kiesling and Wilson (2007) follow an experimental approach to investigate effects of AMPs on market prices and investments. Shawhan et al. (2011) likewise make use of an experimental setting to test the impacts of AMPs and find that firms can influence marginal cost estimates, and thus mitigation measures, strategically. For the suggested best-response approach, we additionally draw from the literature on strategic bidding in multi-unit auctions (e.g. Wolfram, 1999, Wolak, 2003a,b, 2007, Hortaçsu and Puller, 2008) and the literature on the impacts of forward contracts and vertical integration on optimal pricing strategies (e.g. Allaz and Vila, 1993, Wolak, 2007, Bushnell et al., 2008)

The remainder is organized as follows. Section 2 gives an overview of AMPs in US power markets. In Section 3, we outlay and develop the suggested estimation approaches and their empirical implementation. In Section 4, we present the market environment in the Iberian electricity market. Section 5 provides a description of the employed data. In Section 6, we present our results and Section 7 concludes.

2. Automated Market Power Mitigation in US Markets

2.1. Overview and Procedure

Multiple Independent System Operators (ISO) have implemented automated mechanisms for the mitigation of market power exertion in wholesale auction markets. These ISOs are the California Independent System Operator (CAISO), the Independent System Operator New England (ISO-NE), the New York Independent System Operator (NYISO), the Pennsylvania-New Jersey-Maryland Interconnection (PJM), serving various Eastern states, and the Midcontinent Independent System Operator (MISO), whose network also covers parts of Canada. The CAISO, ISO-NE, NYISO and MISO use market observations such as historical bids and prices to construct so called reference levels. Reference levels serve as unit-specific proxies for marginal cost and simulate a competitive bid. The precise derivation of reference levels is further described below. We exclude the PJM, where reference levels are derived by a cost-based method, from our further review. The ISOs are regulated by the US Federal Energy Regulatory Commission (FERC) and publish their full tariffs online, which serve as business practices manuals and operating rules. These FERC-approved tariffs allow an extensive understanding of the procedures applied for automated mitigation, whose generalized concept can be summarized as follows (see Table 1 for an overview).

The basic condition for mitigation is a market situation that implies potential for market power. This is defined by the ISOs as the occurrence of local transmission constraints or as the occurrence of pivotal supply; or both cumulatively. For the latter, a pivotal supplier test is carried out after bid submission that either tests individual suppliers or the group of n-largest suppliers for pivotal supply conditions (MISO, 2019, ISO-NE, 2020, NYISO, 2020). In the case of the CAISO, this screening is further specified by an Residual Supply Index (RSI) analysis (CAISO, 2019).

If this structural test identifies a situation in which there is potential for market power, then respective suppliers' bids are tested against a conduct threshold in order to identify actual exercise of market power. In the case of the CAISO the conduct threshold is met when bids exceed the competitive locational marginal price (LMP) (CAISO, 2019). The other ISOs specify a certain percentage (e.g. 200% or 300%) or absolute amount (e.g. 100$/MWh) by which the submitted bid has to exceed the unit's reference level. If the conduct threshold is exceeded, the bid is deemed non-competitive (MISO, 2019, ISO-NE, 2020, NYISO, 2020).

However, to avoid excess intervention, the bids are then tried against an impact test, which describes the consequential price impact. One possibility is to define the impact as significant as soon as the bid sets the LMP or if the bid effectively removes the unit from the economic merit order (CAISO, 2019). Another possibility is to set an impact threshold as a percentage (e.g. 200%, less for constrained areas) or absolute amount (e.g. 100$/ MWh, less for constrained areas) by which the clearing price would be decreased in a mitigated scenario. This may also be measured by comparing the unit's node's LMP against the node's hub LMP (MISO, 2019, ISO-NE, 2020, NYISO, 2020).

Provided the impact threshold is exceeded, the automated mitigation takes place by overriding the respective bid by a unit-specific reference level. For all analyzed ISOs this practice is applied in day-ahead markets and other spot markets (CAISO, 2019, MISO, 2019, ISO-NE, 2020, NYISO, 2020). Yet, ISOs are heterogeneous when it comes to the possibilities for the calculation of reference levels. The applicability ranking of the available methods is either at the supplier's choice or set by the ISO. The cumulated variety of methods found in the operating procedures of the analyzed ISOs consists of accepted offer-based, LMP-based, and cost-based calculations as well as a negotiation-based method.

The first calculation method is based on previously accepted offer bids of the respective unit and is applied by ISO-NE, MISO and NYISO. In general, the reference level is calculated as the mean or median of accepted offers over the last 90 days during competitive periods, adjusted for changes in fuel prices (MISO, 2019, ISO-NE, 2020, NYISO, 2020).

The second calculation method is based on previous LMPs at the unit's node and is used by all four ISOs. The reference level is calculated as the mean or median of the lowest 25% (50% for NYISO) of LMPs during hours, in which the respective unit was scheduled within the past 90 days. The calculation again includes an adjustment for changes in fuel prices. CAISO additionally distinguishes peak and off-peak hours in the calculation (CAISO, 2019, MISO, 2019, ISO-NE, 2020, NYISO, 2020).

The third calculation method is based on cost estimates and is also applied by all ISOs. This approach considers unit-specific heat rates and fuel cost, unit-specific emissions with respective permit prices, opportunity costs and variable operation and maintenance (O&M) costs. The calculation is done in a consultative approach together with the supplier, who has to provide the required information and documentation of all cost components that cannot be gathered by the ISO (CAISO, 2019, MISO, 2019, ISO-NE, 2020, NYISO, 2020). This approach delivers good estimates of firms' marginal cost, yet requires detailed plant level information on cost structures. Furthermore, regulators are unable to verify the accuracy of data disclosed. Generators naturally have an incentive to overstate their costs, e.g. by overstating the heat rate or the operation and maintenance cost of the power plant.

The last method is based on negotiations and exclusively applied by the CAISO. In this approach suppliers propose an appropriate reference level, which, if not immediately accepted by CAISO, will be further negotiated (CAISO, 2019).

Table 1: Overview of automated market power mitigation across US markets

Procedures	CAISO	ISO-NE	MISO	NYISO
Application tied to transmission constraint	Yes	No	Yes	No
Test for pivotal supply	Yes + RSI	Yes	Partly	Partly
Conduct threshold	Bids exceeding the competitive LMP	% / $ amount per MWh	% / $ amount per MWh	% / $ amount per MWh
Impact threshold	Bid sets LMP/ moves unit out of economic MO	% / $ amount per MWh	% / $ amount per MWh	% / $ amount per MWh
Basis for reference level	a) Prev. LMP b) Negotiated c) Cost	a) Accepted bids b) Prev. LMP c) Cost	a) Accepted bids b) Prev. LMP c) Cost	a) Accepted bids b) Prev. LMP c) Cost
Types of reference levels	Incremental + dynamic cost components	Incremental + dynamic cost components	Incremental + dynamic cost components	Incremental + dynamic cost components
Relevance for day-ahead	Yes	Yes	Yes	Yes

Notes: Summary of the application procedures of automated market power mitigation by different US ISOs. Compiled from CAISO (2019), MISO (2019), ISO-NE (2020), NYISO (2020)

2.2. Calculation of Reference Levels

Our analysis focuses on the estimation of reference levels, which are crucial for efficient mitigation. As the accepted offer-based method is the default method applied by ISO-NE, MISO and NYISO, we use this method as our benchmark. The accepted offer-based method uses previously accepted bids from competitive periods over the recent 90 days as a basis for a mean or median calculation. The definition of competitive periods is, however, not consistent across analyzed ISOs. For the ISO-NE ”competitive” refers to the mere economic scheduling of a unit (ISO-NE, 2020), whereas for the MISO the term is tied to the absence of transmission constraints (MISO, 2019). The NYISO tariff, despite stating the term, does not provide an explicit definition at all (ISO-NE, 2020).

Some ISOs impose additional conditions that narrow down the scope of relevant offers to certain periods or hours within the competitive periods (see Table 2). The NYISO takes only hours into account that start from 6am to 9pm and categorically excludes weekend and holiday hours from the calculation (ISO-NE, 2020). This can be interpreted as an on-peak-focused approach. The MISO does not restrict the calculation to certain hours of the day but instead distinguishes between on-peak and off-peak hours (MISO, 2019). Last, the ISO-NE does not further narrow down the scope of considered accepted bids apart from its definition of competitive periods (ISO-NE, 2020).

Table 2: Conditions for the consideration of previously accepted bids for reference level calculation

Criterion	**ISO-NE**	**MISO**	**NYISO**
Retrospective time frame	90 days	90 days	90 days
Definition of competitive period	Scheduling of the unit in economic merit order	Absence of transmission constraints	None given
Distinction/ exclusion conditions	None given	Distinction of peak and off-peak hours	Only hours starting 6am-9pm, Exclusion of weekends + holidays, Exclusion of bids below 15$/MWh

Compiled from MISO (2019), ISO-NE (2020), NYISO (2020)

The detailed calculation approaches for the default accepted offer-based method reveal a lacking consistency in the definition of which categories of hourly bids are most appropriate as a basis for reference level calculation. From the calculation practices no consensus can be found particularly on the handling of peak and off-peak periods in terms of their distinctive use, inclusion or exclusion. In case of the ISO-NE no attempt of distinguishing peak and off-peak hours is even made, which leads to a rudimentary mean or median calculation. The different approaches to accepted offer-based calculation among the ISOs imply differing calculation results. It is however unclear, which ISO's approach yields reference levels that best approximate competitive bids. Moreover, under certain conditions the ISOs may switch to a cost-based calculation for individual bids. The cost-based methodologies are more uniform among all ISOs as compared to the accepted offer-based methodologies. As a consequence, the cost-based calculation can be expected to yield more similar reference level results across the ISOs, when compared to results from accepted offer-based calculations. This inevitably raises the question of how comparable reference levels of the same ISO really are, if, within the same territory, some bids are regulated using cost-based reference levels, whereas others are regulated using accepted offer-based reference levels.

Both the accepted offer-based calculation as well as the cost-based calculation bear risks of Principal-Agent problems arising from hidden information. As the ISOs rely severely on the accepted offer-based method, this has evoked discussions on possible strategic bidding behavior that aims at increasing reference levels. Shawhan et al. (2011) find evidence in an experimental study that, in case of sufficiently high market power, bidders have an incentive to strategically raise their bids during unmitigated periods and thus manipulate the calculation basis for reference levels – so called reference creep. Currently, this issue is addressed in none of the analyzed ISO tariffs; consequently, there are no measures in place to detect or account for reference creep. The second problem of hidden information arises in the cost-based reference method, where the ISOs depend on suppliers to truthfully disclose information on cost components, which cannot be obtained otherwise by the ISO. This information includes e.g. unit-specific opportunity costs. Depending on the agent to disclose such private, unobservable information provides opportunity for strategic behavior. Even at the PJM, an ISO that is particularly experienced in working with cost-based reference levels, these information asymmetries are hitherto unaddressed. The PJM's independent market monitor describes the occurrence of resulting strategic behavior of market participants in the submission of cost components and criticizes that true competitive proxies cannot be obtained if suppliers' submissions are not truthful and uniform (Monitoring Analytics, 2019). The complexity of bottom-up cost calculation as well as the information asymmetries of this approach may be a reason why all analyzed ISOs, except for the CAISO, explicitly present the cost-based method as least applicable option to calculate reference levels.

3. Method and Empirical Strategy

In this section, we present and develop different empirical approaches to calculate reference levels of power plants' marginal cost based on observed supply bids. To ensure comparability, all approaches make use of the same data from the Iberian day-ahead market, as described below in section 5. First, we present the benchmark procedure as conducted by the NYISO, where we use observations of the preceding 90 days to calculate reference levels. We then proceed by describing the best-response approach, which builds on Wolak (2003a, 2007) and Hortaçsu and Puller (2008). We present two more approaches which are bid pattern-driven and represent extensions to the NYISO benchmark method. Here, we address problems which arise due to start-up cost and reference creep and increase the precision of estimation.

3.1. The NYISO Benchmark Approach

To assess the relative performance of our proposed calculation approaches we first define a benchmark. To that end we choose the NYISO method of calculating reference levels of plants' marginal cost. As compared to other ISOs, the NYISO provides relatively more information on the composition of the calculation basis, i.e. the set of historical bids which is employed for the estimation of reference levels. All US system operators in our analysis follow similar procedures, yet approaches differ in details such as the exclusion of bids from the calculation basis (see Table 2 for an overview).

We calculate reference levels of plants' marginal cost for an exemplary week in December

2017 (4th of December to 10th of December). For each fossil power plant and day within this week, we determine a reference level, which should optimally reflect the bottom-up calculated marginal cost for the respective plant and day.[2] As calculation basis, we use historical bids of the plant within the last 90 days. In line with the NYISO procedure, we define the reference level as the mean or median (whichever is lower) of bids in the calculation basis. Note that we only use bids within the range of 20 €/MWh to 125 €/MWh, firstly to comply with the NYISO procedure, and secondly to limit the leverage of complementary cost considerations of the firms.[3]

Within the 90 days period that serves as calculation basis, variation in underlying fuel cost and cost for carbon emissions is substantial (see Table 3). The precision of reference levels on the one hand benefits from the large calculation basis, but should, on the other hand, not be affected by changes of input prices. System operators account for fuel price changes NYISO (2020), yet do not specify how they proceed exactly.[4] We present our strategy to empirically control for changes in input prices in the Appendix.[5] Reference levels are then defined as the mean or median of all adjusted bids in competitive hours within the last 90 days.

3.2. Best-Response Bidding

The second approach is based on Wolak (2003a, 2007), who derives underlying marginal cost directly from observed bids. We use his model of best-response pricing, which assumes according to supply function equilibria (Klemperer and Meyer, 1989) that a profit maximizing

[2]We present a detailed description of our bottom-up calculation of "true" marginal cost is section 5.

[3]Companies alienate simple bids to signal that a plant is already running (by bidding at very low prices), or that it would need to start-up (by bidding close to the price cap) (Reguant, 2014).

[4]Adjustments are contingent on detailed price information over time. As fuel prices and emission allowance prices are publicly available, we assume that regulators possess the required information.

[5]This input price adjustment does not only include fuel prices but also emissions allowance cost, following Fabra and Reguant (2014), who show that emission cost are passed through at high rates.

firm will submit a set of bids that is ex-post optimal given its residual demand. Assuming profit-maximizing behavior, it is possible to derive a firm's marginal cost C' for observed residual demand RD, observed market clearing prices p and its forward contracted quantity QC.[6] The resulting firm profit function for a single scheduling hour is further dependent on the price received on forward sales PC as well as the uncertain demand shock η and can be expressed as follows:

$$\pi(p) = RD(p, \eta)p - C(RD(p, \eta)) - (p - PC)QC, \tag{1}$$

We take the first order derivative with respect to the price and solve for the marginal cost component to receive the following condition:

$$C'(RD(p^*, \eta)) = p^* - \frac{QC - RD(p^*, \eta)}{RD'(p^*, \eta)} \tag{2}$$

All bids are submitted in the expectation that the respective bid could determine the market clearing price, therefore each bid can be regarded as an optimal price p^*. Marginal cost C' are thus derived from observed bid levels p^*, the amount of infra-marginal quantity offered by the firm RD, the slope of the residual demand function faced by the firm RD', and its contracted quantity QC.[7] As we possess information on all supply and demand bids as well as the owning structure of the firms, we can derive the infra-marginal quantity and the residual demand curves. However, residual demand functions are step-wise bid functions in electricity markets and not continuously differentiable. We follow Wolak (2003a) and solve

[6] Bohland and Schwenen (2022) use a similar framework to analyze the effect of renewable subsidies on strategic pricing.

[7] Firms owning a larger portfolio can strategically play on this portfolio (and potentially market power), leading to a supply function whose underlying true marginal cost might not be non-decreasing. This implies the the marginal cost derived by the best-response bidding model might calculate a marginal (opportunity) cost at the firm and not the unit level. Using this marginal cost as a unit-specific reference level in mitigation however could incentivize firms to bid truthfully according to non-decreasing actual unit marginal cost to avoid disadvantageous reference levels.

this by applying smoothing parameters for the residual demand curve.[8]

The contracted quantity QC is a crucial element for the bidding strategy of the firm. It incorporates both, forward sales (Wolak, 2007, Holmberg, 2011) as well as resell obligations of vertically integrated retailers (Kühn and Machado, 2004, Mansur, 2007, Bushnell et al., 2008), as the underlying incentives are identical. If the contracted quantity exceeds sales in the market, the firm acts as a net-buyer and aims at lowering the market clearing price by bidding below marginal cost. If market sales exceed the contracted quantity, the firm acts as a net-seller and bids above marginal cost to increase its profits. In case the regulator possesses information on vertical sales and forward contracts, it can directly derive QC and thus the underlying marginal cost C'. Unfortunately, we lack information on firms' forward sales and need an alternative approach for the estimation of QC. We make use of the nature of firm strategies and identify the contracted quantity as the position where the marginal cost curve of a firm intersects its supply function (Hortaçsu and Puller, 2008). The rationale is that if the uncertain residual demand materializes at the exact contract position of the firm, the firm has no incentive to influence the market clearing price and bids equal to marginal cost.[9]

We derive all parameters of equation 2 and calculate marginal cost as a function of the observed bid-level, the firm's hourly net-position, and the slope of the residual demand curve at the chosen bid-level. We determine reference levels for all fossil plants and days within a week in December 2017 (4th of December to 10th of December). To ensure comparability across methods, we again restrict input bids to the range from 20 €/MWh to 125 €/MWh in competitive hours (from 7am to 11pm). Last, we define daily reference levels for each

[8]We use the *monpol* function in R, which is part of the *MonoPoly* package and ensures a monotonic fit. We allow for nine degrees of freedom. Note that our findings are not contingent on the exact specification of smoothing parameters.

[9]To retrieve the intersection between the supply curve and the marginal cost curve, we first need to fit a marginal cost curve. We use an isotonic regression that delivers monotonically increasing step-functions and is best-suited to mimic the nature of marginal cost curves.

plant as the mean of all calculated marginal cost estimates for the respective plant and day.

3.3. Accounting for Start-up Cost

In this section we present an extension of the benchmark NYISO method. By following the NYISO approach as presented in section 3.1, we do not structurally incorporate additional cost components such as start-up cost. Yet, the bids in our calculation basis may partly be driven by the presence of start-up cost. Reguant (2014) shows that the neglect of start-up cost leads to biased estimates of marginal cost and eventually to flawed mark-ups and measures of market power. Nevertheless, for the sake of simplicity and clarity, we abstain from including start-up cost in the bottom-up calculated marginal cost estimates.[10] We assess the performance of the presented approaches by the deviation between the respective reference levels and the bottom-up estimates of marginal cost. To achieve coherence, we thus need a calculation basis that excludes bids driven by start-up cost.[11]

Empirically, we address this problem by further limiting our calculation basis to those plants which are clearly not affected by start-up cost. Firms submit very low first step bids for plants that are already running to ensure that these plants will be called with certainty (Reguant, 2014). Note that firms are permitted to submit up to 25 discrete steps per power plant. Using the first step to determine whether the plant should be running or not therefore comes at negligible opportunity cost. We make use of this signaling behavior and limit the calculation basis to bids of power plants for which at least one low-priced bid has been submitted within the respective hour.[12] Apart from this constraint, we use the

[10] A distinct assessment of start-up cost is difficult as in some cases firms make use of complex bids to express start-up cost, whereas in other cases they incorporate them in simple bids.

[11] The alternative would be to include start-up cost in the bottom-up estimates of marginal cost and in the reference levels. However, we see no feasible option to determine the extent to which a bid is driven by start-up cost.

[12] We set the boundary at 30 €/MWh and thus significantly below the minimum clearing price within our sample period, which equals 41.1 €/MWh and which is also below bottom-up estimates of marginal cost as seen in Figure 3.

same calculation basis as in our benchmark approach (see section 3.1) and likewise account for changes in input prices.

3.4. Clustering

In our final approach, we address several additional shortcomings of the NYISO method, namely the large dispersion of results across power plants, the missing calculation basis for a set of plants,[13] and the potential occurrence of reference creep. We tackle these problems by departing from the calculation of unit-specific reference levels. Instead, we apply a machine learning algorithm (k-means clustering) to cluster the 89 power plants in our sample with respect to their main characteristics relevant for marginal cost, i.e. efficiency (serving also as a simultaneous distinction by fuel type) and size. Figure 1 depicts the results of the clustering process, showing four clearly distinguishable clusters. Clusters one and two incorporate large (cluster 1) and small (cluster 2) coal power plants, whereas clusters three and four show large (cluster 3) and small (cluster 4) combined-cycle gas turbines (CCGT).

We use these clusters and calculate reference levels analogously to our procedure in section 3.3, yet not for each power plant individually, but at the cluster-level. Thereby we solve the problem of the large dispersion of estimation errors across plants and receive a more concentrated distribution of results. At the same time we limit the influence of outliers, which are usually attributed to a small calculation basis or market power abuse. Furthermore we solve the problem of missing calculation bases. As the calculation basis is now identical for all power plants within a cluster, we obtain reference levels for a larger set of power plants.

For the purpose of AMPs, the main advantage of clustering the plants is the prevention, or at least complication, of reference creep. As long as reference levels for mitigation are

[13]This pertains to plants who had been recently inactive in the market, e.g. due to maintenance, or to new generating units entering the market.

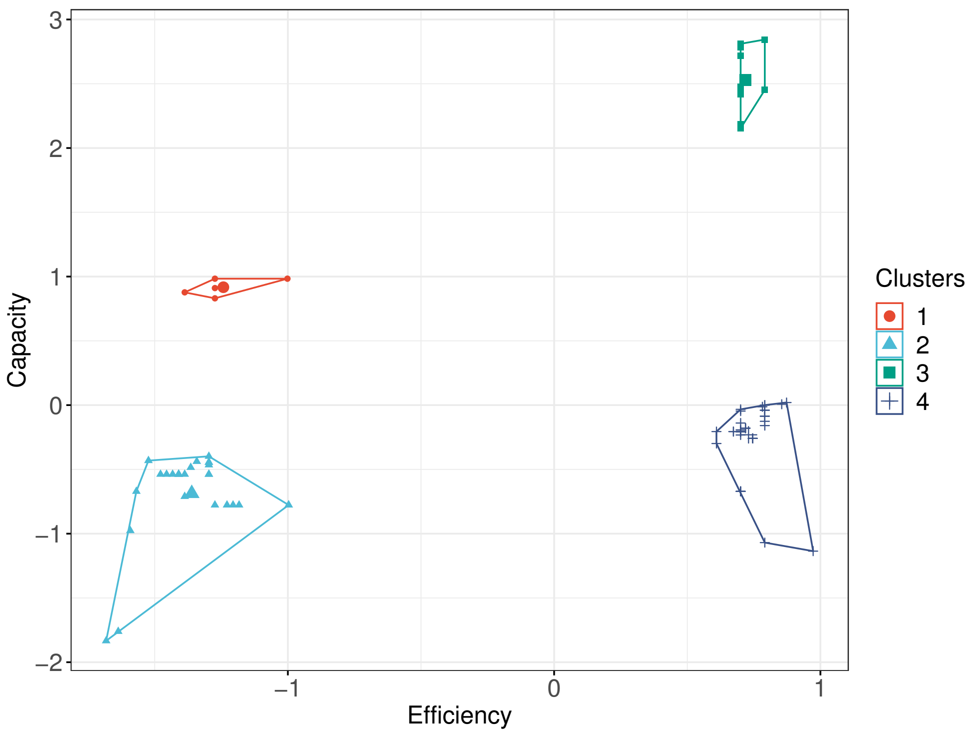


Figure 1: Clustering of the 89 sample plants with respect to efficiency and size. **Clusters 1 and 2** represent inefficient coal power plants, where cluster 1 comprises large coal power plants and cluster 2 small coal power plants. **Clusters 3 and 4** represent efficient CCGT plants with cluster 3 comprising large CCGT plants and cluster 4 smaller CCGT plants. Clustering by efficiency makes additional clustering by fuel-type obsolete, because coal and gas power plants are on different ranges of the efficiency spectrum for technological reasons.

merely based on the historical bids of a single power plant, strategically inflating these bids may prove to be beneficial for the firm. The incentives and ability to strategically alter the calculation basis decrease when the regulator shifts to a clustered approach. Firstly, strategic bidding would become more apparent as the clusters comprise plants of similar size and efficiency. Strong deviations from the mean bidding behavior of the plants within the cluster would be conspicuous and could hardly be justified. Secondly, plants within a cluster belong to a set of different firms as long as clusters are sufficiently large. Conducting reference creep would thus require significant coordination among firms. The clustering approach thus solves and mitigates several elementary problems of accepted offer-based calculations of reference levels.

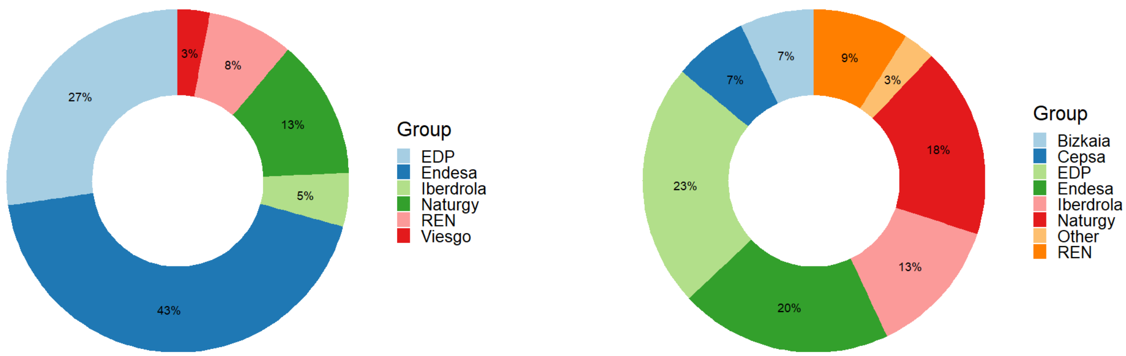
4. Market Environment

The Iberian electricity market consists of the geographical regions of Spain and Portugal. In 2007 the two countries integrated their electricity markets into one administrative market called Mercado Ibérico de la Electricidad (MIBEL). The peninsular electricity spot market of MIBEL is managed by the nominated electricity market operator called Operador del Mercado Ibérico de Energía – Polo Español (OMIE), which is based in Spain. The organized forward market is managed by the Portuguese equivalent OMIP.

OMIE is responsible for the MIBEL day-ahead and intraday (auction and continuous) energy markets within the spot market management. The OMIE market represents the most important place of electricity exchange within MIBEL, as its markets traded 85% of the total MIBEL electricity demand in 2017, which is our year of study. Whenever interconnections between Spain and Portugal are not at capacity limits, OMIE consists of only one pricing zone. This was the case in 94.4% of the time in 2017. The OMIE market can therefore be regarded as one coupled market consisting of the geographic zones of peninsular Spain and Portugal.

This study concentrates on OMIE's day-ahead market, as it represents the most important trading market accounting for more than 86% of the total OMIE trading in 2017. In 2017, a total of 247 TWh was traded in the day-ahead market, of which Spanish generation accounted for the large majority of 72%, whereas Portuguese day-ahead generation accounted for 22%. On the day-ahead market, agents submit supply (sale) and demand (purchase) bids on electricity transactions for the following day. Buying agents can be direct consumers, retailers, resellers and representative agents; selling agents can be owners of production units, retailers, resellers and representative agents (OMIE, 2015).

The daily scheduling horizon consists of 24 hourly periods, which are all auctioned in a single session. Each bid is comprised of up to 25 blocks for each hourly scheduling period,



(a) Coal-fired

(b) Natural gas-fired

Figure 2: Distribution of fossil power generation across firms (5th of September 2017 to 10th of December 2017)

with decreasing prices for demand bids and increasing prices for supply bids. The maximum possible bid price is regulated to 180.30 €/MWh. Demand bids are always simple bids, meaning that they consist only of a price and an amount of power for each block of a scheduling period. Supply bids are tied to a production unit and can be either simple (only price and amount) or complex. Complex bids contain additional conditions that the agent can submit to the market operator and typically cover complementary cost factors such as start-up or ramping cost. OMIE verifies the bids and matches supply to demand bids with the Euphemia matching algorithm that is commonly used in multiple European electricity markets. The algorithm creates two aggregate step-wise curves for demand and supply bids, considering any complex conditions, and finds the corresponding system marginal price as a uniform clearing price (OMIE, 2015).

The day-ahead market is characterized by the presence of few large players dominating the market. Roughly two thirds of generation can be accounted to five company groups owning the respective generation units, namely Endesa, Iberdrola, EDP, Naturgy, and Viesgo (Comisión Nacional de los Mercados y la Competencia, 2019). At the same time, these companies are vertically integrated, and likewise act as electricity resellers and retailers. With small renewable producers entering the market, the overall market share of the dominant producers shrank after liberalization. Yet, the fossil fuel production, which is at the center of our research, is still in the hands of a few large companies. Only six companies accounted for total production from coal-fired units within our sample period, namely Endesa, Iberdrola, EDP, Naturgy, Viesgo and REN. Production from natural gas-fired CCGTs stemmed from the same companies along with Engie, Cepsa, and Bizkaia. These seven companies were responsible for 97 % of natural gas-fired production within our sample period. Figure 2 visualizes the highly concentrated market environment of fossil power production in Spain and Portugal.

5. Data

The centerpiece of our dataset stems from the Iberian market operator OMIE and comprises all supply and demand side bids in the Iberian day-ahead market.[14] Our analysis focuses on fossil power generation, i.e. power production from coal and natural gas. Therefore, we chose a sample period with a high market share of fossil production. The week we analyze in detail is week 49 in 2017, starting on December 4th and ending on December 10th. As we need input data that stretches back 90 days, our sample includes all bids from 5th of September to 10th of December and extends over a period of slightly more than three months.

We focus on fossil production as we compare the derived reference levels to bottom-up calculated marginal cost. For fossil generation this calculation is straight forward and delivers precise estimates of the true underlying marginal cost.[15] Our bottom-up estimation of short-run marginal cost includes fuel cost, cost for carbon emissions, variable O&M cost as

[14]Monthly files including all supply and demand curves are provided online.

[15]Nuclear generation as must-run generation is usually always bid into the market at low prices and therefore market power issues do not play a relevant role. Renewable generation is also bid into the market at low cost due to marginal cost being virtually zero. We further exclude hydro power as hydro bids represent the dynamic value of water, which is strongly driven by opportunity cost.

well as all relevant additional taxes and levies. Figure 3 displays the estimated marginal cost across both technologies in Spain and Portugal. For a detailed overview of the determinants of our calculation, as well as sources of fuel prices and plants' efficiency rates, please see Table A.1 in the Appendix. Table A.2, likewise displayed in the Appendix, provides the detailed magnitudes of parameters we use for our calculation.

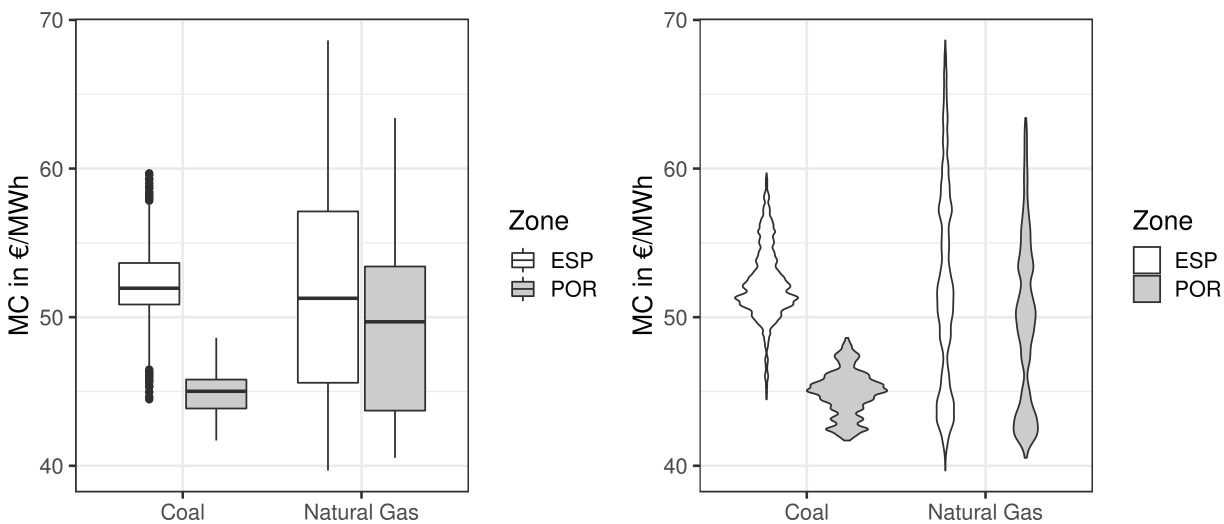


Figure 3: Distribution of bottom-up estimated marginal cost across fossil power plants within the sample period (5th of September 2017 to 10th of December 2017)

Figure 3 gives an overview of the bottom-up estimated marginal cost in our final sample. Note that taxes and levies in both country jurisdictions structurally differ, attributing to the structural marginal cost difference between Spanish and Portuguese plants. The initial reason stems from the additional taxation prevalent in Spain. Even though Portugal implemented a clawback mechanism to mitigate the difference in marginal cost via an additional fixed charge, this mechanism lacks the ability to fully compensate the cost gap. At the same time it is apparent that marginal cost of coal power plants are subject to less volatility than marginal cost of CCGT plants, which is attributed to the higher volatility of natural gas prices as compared to hard coal prices.

As part of our analysis is based on firm behavior, we additionally assign the parent companies to each power plant, or more precisely, to each bid, to account for ownership structures. This provides us with a dataset that comprises all demand and supply bids within the sample period, enriched by bottom-up estimated marginal cost, information on fuel types, and an indicator variable specifying the owning parent company of the respective plant.

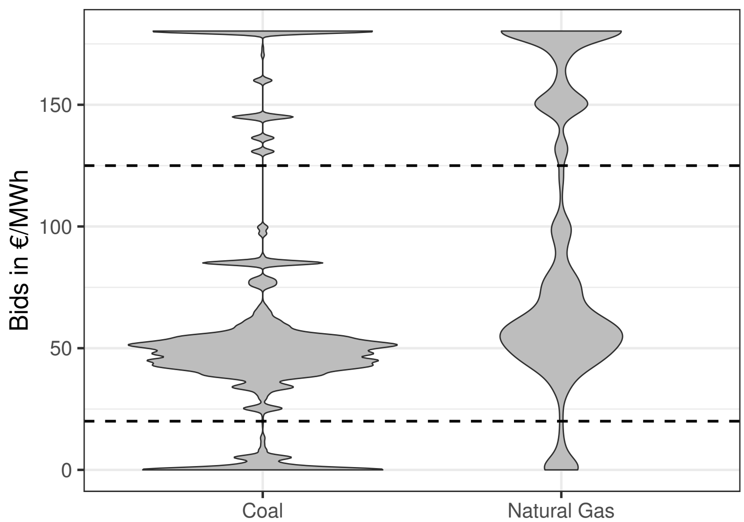


Figure 4: Distribution of bids submitted by fossil power plants within the sample period (5th of September 2017 to 10th of December 2017)

For the benchmark method to calculate reference levels of underlying marginal cost, we mimic the procedure of the NYISO and take it to the Iberian data. Analogous to the NYISO procedure, we thus restrict our calculation basis to a certain range of bids deemed competitive according to the NYISO rationale. In the NYISO calculation, all bids lower than 15 $/MWh are excluded. We apply an analogous boundary at 20 € /MWh and furthermore set an upper boundary of 125 € /MWh to exclude miscellaneous bids. This means we exclude all those bids, which we are sure not to reflect short-run marginal cost but rather assume to be signaling behavior (must-run/ must-not-run). Figure 4 displays the observed bid levels of both technology types within our sample period, as well as the cut-offs at 20 €/MWh and 125 €.[16] Even though firms can make use of complex bids to cover cost complementarities such as start-up or ramping cost, they simultaneously use simple bids to either ensure that the respective power plant is running (and bid close to zero), or to signal that they not intend to start-up a plant (and bid close to the price cap). This explains the high density of bid levels at 0 €/MWh and 180.30 €/MWh as displayed in Figure 4. Additionally, we limit the sample to competitive hours (from 7am to 11pm) on weekdays to be consistent with the NYISO procedure.

In Table 3, we present the summary statistics of our final sample. Note that the dispersion of natural gas prices by far exceeds the dispersion of hard coal prices, further shedding light on the distribution of marginal cost in Figure 3.

6. Results

In this section we present the results of our empirical analysis. We first present results for the different approaches to reference level calculation. Based on this, we apply our preferred approach to a simulation of automated market power mitigation and analyze welfare effects.

6.1. Calculating Reference Levels

As described in detail in section 3, we tested the benchmark approach as well as three alternative approaches to calculate reference levels of marginal cost. We assess the performance of the approaches based on two quality criteria. First, we compare the mean absolute error between the derived reference levels and the true marginal cost. The second criterion for the performance of each estimation method is the number of covered plants. The more we

[16] Out of the total 710,611 fossil oberservations in our sample period this leads to a cut-off of 32.53 %, out of which the majority are natural gas bids.

Table 3: Summary statistics

	Mean	Median	Std. dev.	Min	Max	Obs.
Coal bid level [€/MWh]	50.3	48.7	12.7	22.4	100.0	122,655
Gas bid level [€/MWh]	59.1	55.5	17.1	20.1	123.8	135,239
Coal marginal cost [€/MWh]	50.5	51.2	3.9	41.7	59.7	122,655
Gas marginal cost [€/MWh]	51.9	51.7	6.6	41.0	68.6	135,239
Coal mark-up [€/MWh]	-0.1	-2.2	11.1	-31.2	50.1	122,655
Gas mark-up [€/MWh]	7.1	3.2	17.3	-42.3	78.0	135,239
Coal bid size [MWh]	45.5	36.5	48.1	0.3	555.0	122,655
Gas bid size [MWh]	65.1	30.0	94.1	0.2	805.0	135,239
Clearing price [€/MWh]	61.7	61.5	9.4	41.1	170.0	4160
Hard coal price [€/MWh]	10.7	10.7	0.3	10.1	11.1	69
Natural gas price [€/MWh]	21.9	21.9	3.4	17.1	30.2	69
EUA price [€/ton of CO2]	7.3	7.4	0.3	6.5	7.9	69

Notes: Sample from 5th of September 2017 to 10th of December 2017 for hours 8 to 23, excluding Saturdays and Sundays. Sample is further restricted to bids higher than 20 € /MWh and lower than 125 € /MWh. Observations are hourly and comprise bids from nine large carbon emitting power producers (EDP, Iberdrola, Endesa, Naturgy, Viesgo, REN, Cepsa, Engie, and Bizkaia)

restrict the calculation basis within our empirical setting, the lower the number of plants for which we obtain reference levels. To ensure stable operation of an AMP, reference levels should at best be available for all power plants in the market.

In Table 4, we present our main findings.[17] The benchmark NYISO approach clearly performs worst and exhibits a mean absolute error across plants of 11.52 € /MWh. The best-response approach delivers smaller mean error terms as well as less dispersed outcomes across plants. Moreover, the maximum error term falls short of what we observe for the benchmark approach.

For the start-up approach, where we exclude bids from the calculation basis that could be driven by complementary cost factors, we receive a low mean error of 7.27 € /MWh, which clearly constitutes an improvement over the benchmark method. Yet, the lower error comes

[17]In the Appendix we present a similar table on mean errors in relative terms.

Table 4: Deviation from true marginal cost in absolute terms

Approach	Mean	Median	Std. dev.	Min	Max	Plants
NYISO [€/MWh]	11.52	6.95	14.27	0.19	67.76	82
Best response [€/MWh]	8.92	5.10	7.60	1.26	36.76	85
Start-up [€/MWh]	7.27	4.32	9.55	0.24	61.57	72
Clustering [€/MWh]	2.77	2.06	1.84	0.21	9.50	89

Notes: Deviation is defined as the difference between derived reference levels and the true marginal cost we calculated bottom-up. In total, there are 89 power plants in our sample.

at the price of a reduced set of plants due to the restricted calculation basis.

Our last approach overcomes this downside and delivers reference levels for all 89 fossil power plants in our sample. The clustering approach thus covers the broadest set of power plants, which is a crucial aspect for the potential application in AMPs. At the same time it delivers reference levels that lead to the lowest mean error terms of just 2.77 € /MWh.

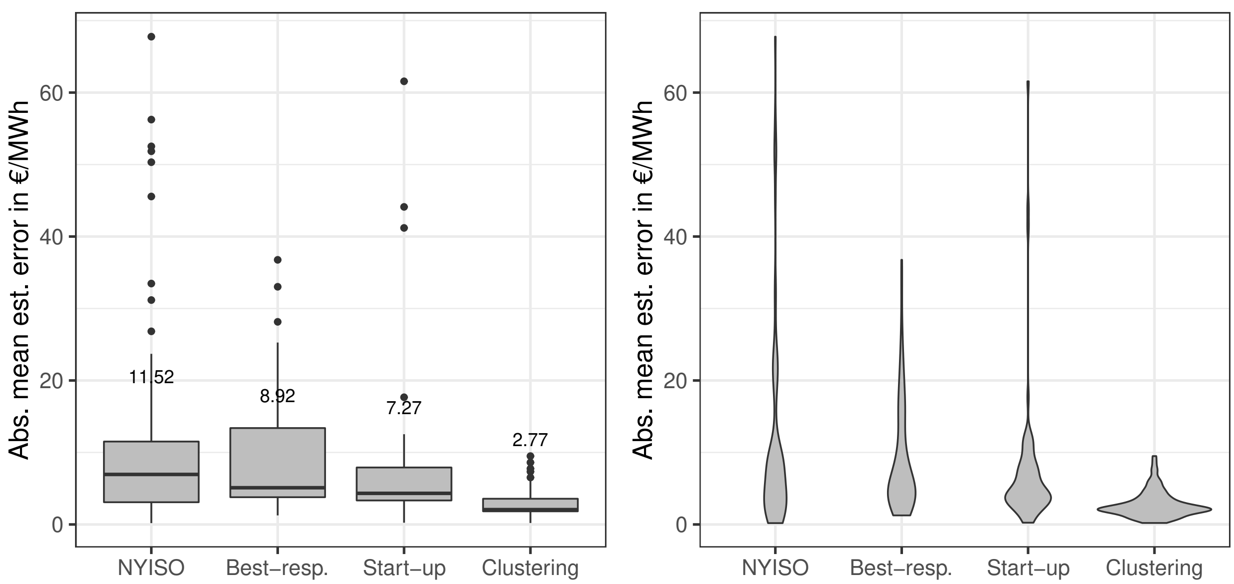


Figure 5: Accuracy of marginal cost estimation across approaches in absolute terms.

The box-plots and violin-plots in Figure 5 and Figure 6 illustrate graphically that all proposed alternatives outperform the method which is currently applied by the NYISO. We deem absolute values of deviations from the underlying marginal cost to be better suited to assess the performance of an approach than relative deviations. Ultimately, a regulator applying automated mitigation or a researcher who seeks to receive appropriate estimates of marginal cost, is mainly interested in achieving precise estimation. Under or overestimation are both undesired.

Nevertheless, it is relevant whether a method leads to systematic positive or negative bias. To that end, Figure 6 shows our results in relative terms.[18] It is apparent that overestimation of marginal cost is more prevalent than underestimation. The preponderance of overestimation is especially pronounced in the NYISO approach and the start-up approach. In an AMP environment, overestimation may turn out to be costly for consumers as incidents of market power exertion could stay unnoticed due to erroneous high reference levels.

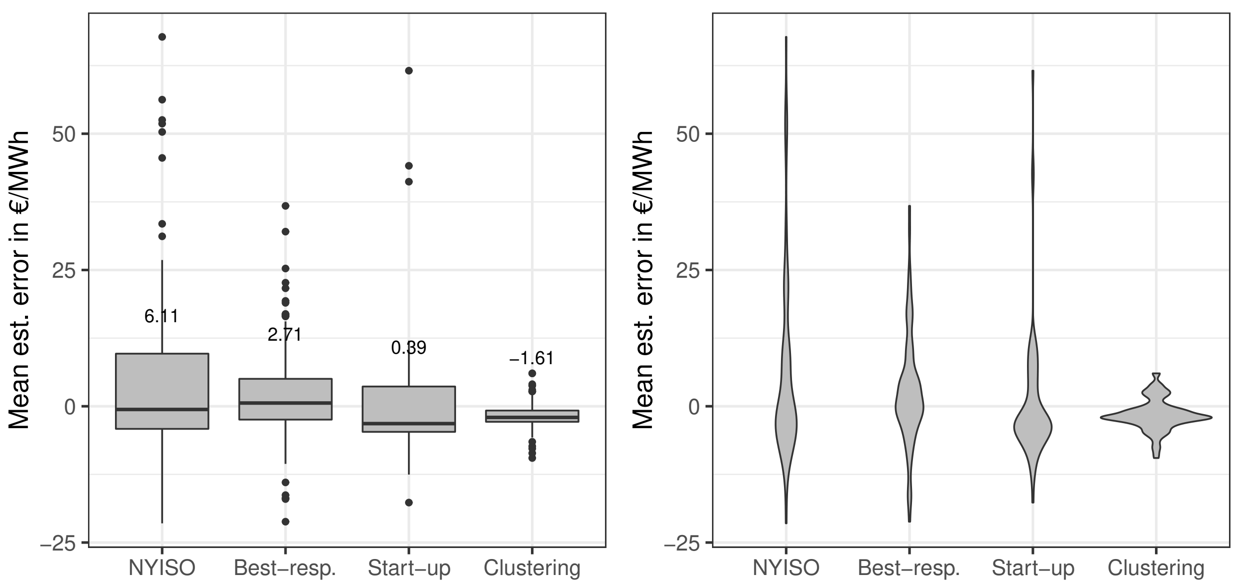


Figure 6: Accuracy of marginal cost estimation across approaches.

Following the model-driven best-response approach leads to more evenly distributed er-

[18] Table A.3 in the Appendix displays the outcomes in more detail.

rors and less outliers (see Figure 6). This approach performs well, but requires additional information on firms' contract positions within the market. Moreover, it can easily be subject to strategic behavior, as firms are able to influence the reference levels in real-time. Among the other three approaches, reference levels are predominantly lower than true marginal cost, indicating a slight structural bias. This bias is driven by coal power plants, for which bid levels often fall short of marginal cost. When firms need to meet certain contract obligations, they often price below marginal cost. As coal power plants are usually situated to the left of CCGT plants within the merit order, coal power plants are more affected by these strategic considerations. If mitigation measures were to be implemented strictly, systematic underestimation of marginal cost would harm producers, as mitigation would enforce bids below true marginal cost. However, this problem is addressed in the conduct test by granting a predefined margin by which bids can exceed the reference level.

6.2. Mitigation Simulation

We have now established the clustering approach as our most preferred way of calculating reference levels due to superiority in precision, coverage and risk reduction of reference creep. In order to quantify welfare impacts that this mitigation mechanism would have on a previously unmitigated market like the Iberian day-ahead, we apply this approach in a simulation of automated mitigation. For our sample estimation week from December 4th to December 10th, we apply the multi-step mitigation procedure outlaid in Section 2.

Conduct test. We submit all bids to a conduct test, which bids fail if they exceed their respective daily reference level by more than a 20 € or 50 % threshold.[19]

Impact test. For hours where bids have failed the conduct test, we perform an impact

[19] These values follow the thresholds from the NYISO benchmark approach as we want to refrain from using arbitrary values. If AMPs were actually implemented it would of course make sense to consider adapting them to the specific market conditions.

test. This test evaluates if, compared to a mitigated market supply curve, the clearing price is increased by more than a 20 € or 50 % threshold.[20] We calculate the counterfactual mitigated clearing price by constructing a new supply curve ("impact test supply curve"). Bids that have passed the conduct test enter this curve at their original level. Bids that have failed the conduct test enter this curve at their reference levels. We then calculate the impact-clearing price by finding the intersection of the original step-wise demand curve and the step-wise impact test supply curve as illustrated in 7. As a last step, we compare the original clearing price with the impact-clearing price to determine if the above impact thresholds were exceeded.

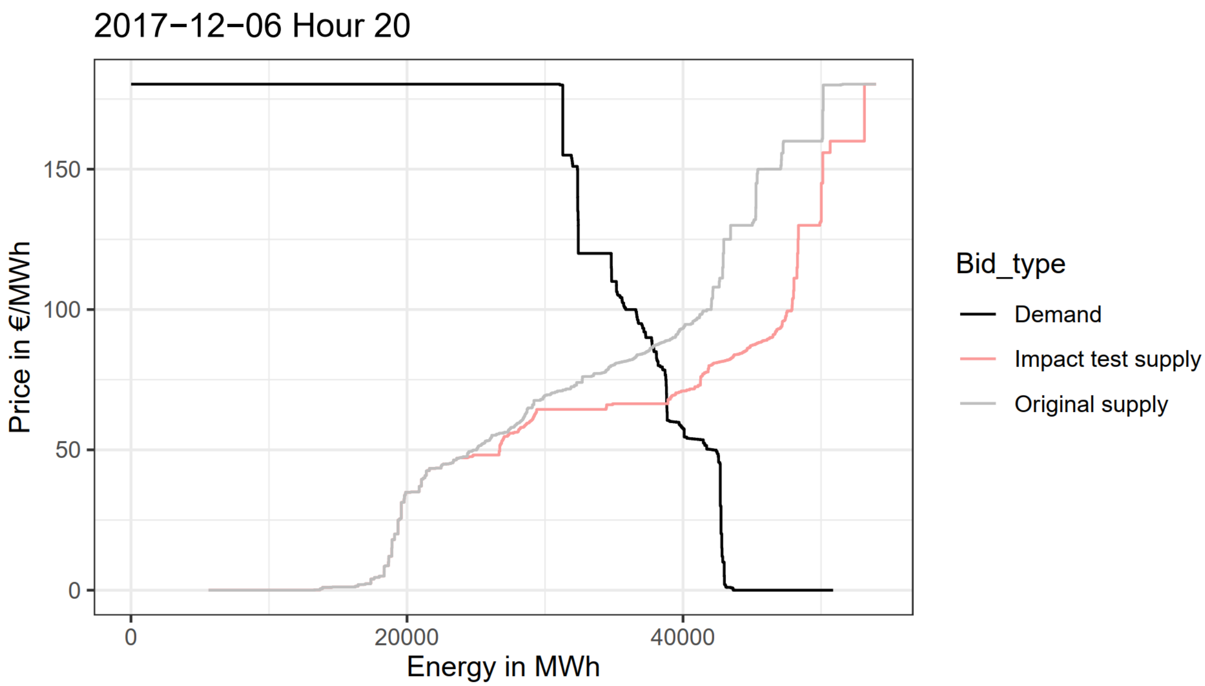


Figure 7: Original and resulting market clearing curves of impact test for an exemplary hour.

Mitigation. In hours, in which both tests fail, we perform actual bid mitigation of conduct-non-conform bids to their respective reference levels. The new clearing price of

[20]As the Iberian day-ahead market does not have nodal or zonal pricing, we perform the impact test against a collectively mitigated scenario of the whole market.

these hours is now the clearing price calculated in the impact test. Out of the 168 hours in our weekly sample, mitigation occurs in 4 hours, which appears as a somewhat reasonable incidence of market interference.

Welfare impacts. For the 4 mitigated hours we find a notable, dead-weightloss-decreasing rise in market efficiency, amounting to 6.57 % increased social welfare for these hours. This goes along with sizeable welfare transfers from supplier surplus to buyer surplus. In the mitigated hours supplier surplus on average decreases by 32.90 % and buyer surplus increases by 25.60 %.

Welfare robustness. We have to consider however that the reference levels, to which non-competitive bids are mitigated, are only a proxy for marginal cost. The true supplier surplus and true welfare impacts, based on true marginal cost, may therefore deviate. In order to calculate the true welfare impacts as a robustness check, we apply the same merit order but instead of taking the (mitigated) bids to calculate supplier surpluses, we take our bottom-up engineering estimates of marginal cost. The resulting true losses in producer surplus are 47.33 % for mitigated hours. The overall impact on true social welfare is slightly lower than the observed one, yet still sizeable at 6.51 %. We can therefore conclude that not only observed welfare would increase thanks to mitigation, but also true welfare would increase by a similar magnitude.

7. Conclusion

This paper contributes to improved automated mitigation of market power in electricity markets. Automated mitigation procedures (AMPs) find wide application in US power markets and are designed for real-time detection and mitigation of market power abuse. AMPs rely on so-called reference levels, supposed to approximate marginal cost, to evaluate competitiveness of a bid and to mitigate it by overriding. We design alternative approaches to derive reference levels from producers' supply offers. Improved accuracy of marginal cost estimates allows for both, facilitated detection of market power, as well as refined and more targeted mitigation. Refined mitigation protects buyers from excessive redistribution of rents to suppliers, but in a given mitigation setting likewise protects suppliers from excessive and unjust mitigation of competitive offers.

We employ micro-level data from the Iberian day-ahead market to test our suggested approaches to deriving reference levels against a best-practise benchmark. As benchmark approach, we choose the procedure as followed by the New York Independent System Operator, where reference levels are inferred from past offers of a power plant. In our application of this benchmark approach, we find deviations of marginal cost estimates from true marginal cost to be substantial, with a mean absolute deviation of 11.52 €/MWh. In comparison, the alternative approaches we propose deliver mean absolute deviations ranging between 2.77 €/MWh for our novel clustering approach and 8.92 €/MWh for the best-response approach based on Wolak (2003a, 2007), where we reverse-engineer marginal cost from real-time hourly offers instead of past offers of a plant. For the clustering approach we depart from the estimation of marginal cost on the unit-level and estimate marginal cost for clusters of similar power plants. This preferred approach of ours does not only yield the most precise estimates, but likewise counteracts reference creep, i.e the strategic manipulation of bids to evade mitigation. System operators should hence consider the adoption of this approach for AMP purposes.

We finally apply our preferred approach in a simulation setting of AMP. We find a mitigation incidence of 4 out of 168 hours, which is associated with notable welfare implications. In mitigated hours buyer surplus increases on average by 25.60 %, supplier surplus decreases by 32.90 - 47.33 %. Overall welfare increases by 6.51 - 6.57 %.

We contribute to potential improvement of policies in electricity markets with market power issues, e.g. related to locational pricing, pivotal supply, and concentrated or integrated market structures. The EU has, for instance, signaled in light of REPowerEU initiatives to reassess locational pricing in the EU and to "ensur[e] an up to date and robust framework to protect against [market power] abuse [...] in periods of high prices and market volatility"[21]. Any applied frameworks will have to make sure (1) that supply bids are fair and competitive and (2) that underlying fluctuations in input prices are taken into account to not harm the profitability of producers. AMPs are a suitable tool to achieve both. The recent power crisis due to the Russo-Ukrainian war is just an extreme example of flexible fossil power generation being the marginal technology and causing high clearing prices with high windfall profits for inframarginal producers. This can be exploited especially by firms who can strategically deploy a technology portfolio. These constellations will continue to occur in decarbonizing electricity systems with increasing shares of cheap, intermittent renewables and limited storage capacities (Graf et al., 2021).

To conclude, we show that current AMPs can be improved considerably by redesigning the estimation of underlying marginal cost of production. This significantly improves market efficiency by means of social welfare increases along with redistribution of excess rents from suppliers to buyers. Moreover, our enhanced approaches facilitate research whenever scholars require cost estimates for empirical analysis in power markets. Our findings are likewise applicable to other use cases and markets, such as monitoring of renewable energy auctions or market power surveillance in air and rail traffic.

[21] https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=COM:2022:236:FIN

References

Allaz, Blaise and Jean-Luc Vila (1993) "Cournot competition, forward markets and efficiency," *Journal of Economic Theory*, 59 (1), 1–16, 10.1006/jeth.1993.1001.

Bohland, Moritz and Sebastian Schwenen (2022) "Renewable support and strategic pricing in electricity markets," *International Journal of Industrial Organization*, 80, 102792, 10.1016/j.ijindorg.2021.102792.

Borenstein, Severin, James B. Bushnell, and Knittel Christopher R.. (1999) "Market Power in Electricity Markets: Beyond Concentration Measures," *The Energy Journal*, 20 (4), 10.5547/ISSN0195-6574-EJ-Vol20-No4-3.

Bushnell, James B., Erin T. Mansur, and Celeste Saravia (2008) "Vertical Arrangements, Market Structure, and Competition: An Analysis of Restructured US Electricity Markets," *American Economic Review*, 98 (1), 237–266, 10.1257/aer.98.1.237.

CAISO (2019) "California Independent System Operator Corporation Fifth Replacement FERC Electric TariffTariff, (Open Access Transmission Tariff), effective as of 28.09.2019," `http://www.caiso.com/Documents/Conformed-Tariff-asof-Sept28-2019.pdf`, accessed on 13.03.2020.

Ciarreta, Aitor and María Paz Espinosa (2010) "Market power in the Spanish electricity auction," *Journal of Regulatory Economics*, 37 (1), 42–69, 10.1007/s11149-009-9102-7.

Comisión Nacional de Energía (2013) "Spanish Regulator's National Report to the European Commission 2013," `https://www.cnmc.es/sites/default/files/1797558_8.pdf`, accessed on 25.05.2020.

Comisión Nacional de los Mercados y la Competencia (2019) "Informe de Supervisiòn del

Mercado Peninsular Mayorista al Contado de Electricidad Año 2017," https://www.cnmc.es/sites/default/files/2322298_0.pdf, accessed on 23.04.2020.

Davis, Lucas W. and Catherine Wolfram (2012) "Deregulation, Consolidation, and Efficiency: Evidence from US Nuclear Power," *American Economic Journal: Applied Economics*, 4 (4), 194–225, 10.1257/app.4.4.194.

EDP (2018) "Annual Report 2017," https://www.edp.com/sites/default/files/annual_report_edp_2017_with_minutes_.pdf, accessed on 25.05.2020.

Fabra, Natalia and Mar Reguant (2014) "Pass-Through of Emissions Costs in Electricity Markets," *American Economic Review*, 104 (9), 2872–2899, 10.1257/aer.104.9.2872.

de Frutos, María-Ángeles and Natalia Fabra (2012) "How to allocate forward contracts: The case of electricity markets," *European Economic Review*, 56 (3), 451–469, 10.1016/j.euroecorev.2011.11.005.

García, José A. and James D. Reitzes (2007) "International Perspectives on Electricity Market Monitoring and Market Power Mitigation," *Review of Network Economics*, 6 (3), 10.2202/1446-9022.1127.

Graf, Christoph, Emilio La Pera, Federico Quaglia, and Frank A. Wolak (2021) "Market Power Mitigation Mechanisms for Wholesale Electricity Markets: Status Quo and Challenges," *Program on Energy and Sustainable Development Working Paper*, https://fsi.stanford.edu/publication/market-power-mitigation-mechanisms-wholesale-electricity-markets-status-quo-and.

Green, Richard (1996) "Increasing Competition in the British Electricity Spot Market," *The Journal of Industrial Economics*, 44 (2), 205, 10.2307/2950646.

Green, Richard J. and David M. Newbery (1992) "Competition in the British Electricity Spot Market," *Journal of Political Economy*, 100 (5), 929–953, 10.1086/261846.

Helman, Udi (2006) "Market power monitoring and mitigation in the US wholesale power markets," *Energy*, 31 (6-7), 877–904, 10.1016/j.energy.2005.05.011.

Holmberg, Pär (2011) "Strategic Forward Contracting in the Wholesale Electricity Market," *The Energy Journal*, 32 (1), 10.5547/ISSN0195-6574-EJ-Vol32-No1-7.

Hortaçsu, Ali and Steven L. Puller (2008) "Understanding strategic bidding in multi-unit auctions: a case study of the Texas electricity spot market," *The RAND Journal of Economics*, 39 (1), 86–114, 10.1111/j.0741-6261.2008.00005.x.

IEA/NEA (2015) "Projected Costs of Generating Electricity 2015," 10.1787/cost_electricity-2015-en.

ISO-NE (2020) "ISO New England Inc. Transmission, Markets, and Services Tariff," `https://www.iso-ne.com/static-assets/documents/regulatory/tariff/sect_3/mr1_append_a.pdf`, accessed on 12.03.2020.

Kiesling, Lynne and Bart J. Wilson (2007) "An experimental analysis of the effects of automated mitigation procedures on investment and prices in wholesale electricity markets," *Journal of Regulatory Economics*, 31 (3), 313–334, 10.1007/s11149-006-9018-4.

Klemperer, Paul D. and Margaret A. Meyer (1989) "Supply function equilibria in oligopoly under uncertainty," *Econometrica: Journal of the Econometric Society*, 1243–1277, 10.2307/1913707.

Kühn, Kai-Uwe and Matilde Machado (2004) "Bilateral Market Power and Vertical Integration in the Spanish Electricity Spot Market," `https://ssrn.com/abstract=608249`.

Mansur, Erin T. (2007) "Upstream competition and vertical integration in electricity markets," *The Journal of Law and Economics*, 50 (1), 125–156, 10.1086/508309.

MISO (2019) "Business Practices Manual – Market Monitoring and Mitigation, Manual No. 009-r15," https://cdn.misoenergy.org//BPM%20009%20-%20Market%20Monitoring%20and%20Mitigation49600.zip, accessed on 12.03.2020.

Monitoring Analytics (2019) "State of the Market Report for PJM," https://www.monitoringanalytics.com/reports/PJM_State_of_the_Market/2019/2019q2-som-pjm-sec1.pdf, accessed on 14.05.2020.

Newbery, David M. (1997) "Privatisation and liberalisation of network utilities," *European Economic Review*, 41 (3-5), 357–383, 10.1016/S0014-2921(97)00010-X.

Newbery, David M. and Michael G. Pollitt (1997) "The Restructuring and Privatisation of Britain's CEGB-Was It Worth It?" *The Journal of Industrial Economics*, 45 (3), 269–303, 10.1111/1467-6451.00049.

NYISO (2020) "NYISO Market Administration and Control Area Services Tariff," https://nyisoviewer.etariff.biz/ViewerDocLibrary/MasterTariffs/9FullTariffNYISOMST.pdf, accessed on 12.03.2020.

OMIE (2015) "Daily And Intraday Electricity Market Operating Rules December 2015," https://www.omie.es/sites/default/files/2019-12/20151223_reglas_mercado_ingles.pdf, accessed on 28.04.2020.

Reguant, Mar (2014) "Complementary bidding mechanisms and startup costs in electricity markets," *The Review of Economic Studies*, 81 (4), 1708–1742, 10.1093/restud/rdu022.

Shawhan, Daniel L., Kent D. Messer, William D. Schulze, and Richard E. Schuler (2011) "An experimental test of automatic mitigation of wholesale electricity prices," *International Journal of Industrial Organization*, 29 (1), 46–53, 10.1016/j.ijindorg.2010.06.005.

Twomey, Paul, Richard J. Green, Karsten Neuhoff, and David Newbery (2006) "A Review of the Monitoring of Market Power The Possible Roles of TSOs in Monitoring for Market Power Issues in Congested Transmission Systems," 10.17863/CAM.5068.

United Nations (2015) "2015 United Nations Energy Statistics Yearbook," https://unstats.un.org/unsd/energy/yearbook/2015/08i.pdf, accessed on 17.06.2020.

Wilson, James F. (2000) "Scarcity, Market Power, and Price Caps in Wholesale Electric Power Markets," *The Electricity Journal*, 13 (9), 33–46, 10.1016/S1040-6190(00)00153-6.

Wolak, Frank A. (2003a) "Identification and Estimation of Cost Functions Using Observed Bid Data," in *Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress*, 2, 133.

——— (2003b) "Measuring unilateral market power in wholesale electricity markets: the California market, 1998-2000," *American Economic Review*, 93 (2), 425–430, 10.1257/000282803321947461.

——— (2007) "Quantifying the supply-side benefits from forward contracting in wholesale electricity markets," *Journal of Applied Econometrics*, 22 (7), 1179–1209, 10.1002/jae.989.

Wolfram, Catherine D. (1999) "Measuring duopoly power in the British electricity spot market," *American Economic Review*, 89 (4), 805–826, 10.1257/aer.89.4.805.

Appendix

Appendix A. Fuel Price Adjustment

The approach for the adjustment of fuel prices is best explained by an example: We want to derive a reference level of marginal cost r for power plant x at a certain day t. This means that for an exemplary bid b within the calculation basis B, submitted at time $t-20$ for power plant x, we can derive a hypothetical efficiency rate ϵ^* that would justify the observed bid level b under the assumption of competitive bidding. Subsequently we use this efficiency rate ϵ^*, as well as current input prices at time t to calculate an adjusted bid b' which becomes part of the adjusted calculation basis B'. Equation A.1 shows the first step, where we equate the past bid b on the LHS with the marginal cost calculation on the RHS.

$$b(x)_{(t-20)} = \frac{Fuelprice_{(t-20)} + CO2price_{(t-20)} * CO2intensity}{\epsilon^*} + O\&M + Taxes\&Levies \quad \text{(A.1)}$$

We solve Equation A.1 for ϵ^*, which captures the level of competitiveness of bid b in $t-20$. We then employ this hypothetical efficiency rate ϵ^* to calculate b' at time t, i.e. the adjusted bid that reflects both, the level of competitiveness of bid b in $t-20$, as well as fuel and emission prices at time t.

$$b'(x)_{(t)} = \frac{Fuelprice_{(t)} + CO2price_{(t)} * CO2intensity}{\epsilon^*} + O\&M + Taxes\&Levies \quad \text{(A.2)}$$

We apply this procedure to each bid in B and end up with the adjusted calculation basis B' that incorporates the competitiveness of bids, net of changes in input prices. From this calculation basis, we then derive the reference level r.

Appendix B. Marginal Cost

Table A.1: Overview of variable cost input data for coal and gas-fired generation

Data type	**Content**	**Scope**	**Source**
Plant efficiencies	Plant-specific efficiency figures where possible; or else average efficiencies acc. to year of commissioning	All coal/ gas-fired plants bid into the day-ahead in 2017	Global Energy Observatory
Coal prices	Daily spot prices for imported coal + RSI	2017	Bloomberg MFE1 COMB
Natural gas prices	Daily spot prices for gas prices in the Iberian gas market	2017	MIBGAS Data 2017, product GDAES_D+1
EUA prices	Daily spot prices for EU-ETS allowances (EUAs)	2017	Bloomberg EEXX03EA
National environmental taxes	1) Taxes on use/ disposal of input resources 2) Energy generation tax (all technologies)	Power plants on Spanish territory; Rate levels of 2017	Ley 15/2012 Título I, Título III; Comisión Nacional de Energía (2013)
Clawback rate	Charge to compensate for unequal tax burdens	Power plants on Portuguese territory; Rate levels of 2017	Decreto-Lei n.º 74/2013 Artigo 1.º; EDP (2018)
Variable O&M costs	Median variable O&M costs per MWh	Coal and gas-fired plants, dataset of 2015	IEA/NEA (2015)

Table A.2: Overview of magnitudes of parameters applied in the marginal cost estimation

Data type	Value	Source
Clawback charge Portugal	6.50 € /MWh until 16.11.2017 4.75 € /MWh as of 17.11.2017	Decreto-Lei n.º 74/2013 Artigo 1.º; EDP (2018)
Energy generation tax Spain	7 % of revenue	Ley 15/2012 Título I
Fossil fuel consumption tax Spain	0.65 € /GJ	Ley 15/2012 Título III
Variable O&M cost coal	2.52 € /MWh	IEA/NEA (2015)
Variable O&M cost gas	3.18 € /MWh	IEA/NEA (2015)
Net calorific value hard coal (averaged for Spain's main import origins Russia, Colombia, Indonesia)	7.333 MWh/t	United Nations (2015)

Table A.3: Deviation from true marginal cost

Approach	Mean	Median	Std. dev.	Min	Max	Plants
NYISO [€/MWh]	6.11	-0.59	17.33	-21.48	67.76	82
Best response [€/MWh]	2.71	0.59	10.24	-21.18	36.76	85
Start-up [€/MWh]	0.39	-3.18	12.03	-17.67	61.57	72
Clustering [€/MWh]	-1.57	-1.99	2.94	-9.61	5.91	89

Notes: Positive values signify that the respective approach delivers higher values than the bottom-up calculation. Deviation is defined as the difference between derived reference levels and the true marginal cost we calculated bottom-up. In total, there are 89 power plants in our sample.