Statistical Arbitrage and Information Flow in an Electricity Balancing Market

Derek W. Bunn[a] *and Stefan O.E. Kermer*[b]

ABSTRACT

Motivated by the events following a natural experiment in 2015, when the market rules for electricity spot trading were changed in Britain, we analyse the operational effects of market participants responding to price incentives for spillage and shortage positions in a single price, real-time market. We develop an analytical model for optimal real-time decisions by generators and speculators based upon forecasts of the conditional distribution of the total system imbalance between instantaneous supply and demand. From this, we examine the effects of time delays in information transparency for the consequent statistical arbitrage positions. We backtested this model empirically to the Austrian system imbalance settlements process within the German/Austrian integrated market. Results suggest that permitting additional intraday flexibility from a physical generator or a non-physical trader can be beneficial for the agents themselves, the system operator and market efficiency.

Keywords: Electricity, Real-time Markets, Trading, Forecasting, Statistical Arbitrage

https://doi.org/10.5547/01956574.42.5.dbun

1. INTRODUCTION

In progressing towards more efficient competitive electricity markets, the liberalizing intent has generally been to replace central control with price signals and markets wherever possible. This is becoming the norm in forward, day ahead, and intraday trading, but in the provision of real-time balancing, progress in this direction has been more cautious. By "real-time" activities we refer to the actual production and consumption of electricity and its instantaneous matching on the network (i.e. "balancing") by the system operator. Prior to this physical delivery to meet demand, all other electricity trading is in forward commitments. Balancing arrangements tend to be idiosyncratic in details, according to system operations and jurisdictions around the world, but a common element is that at some point in time ahead of real-time delivery, forward trading between market participants has to stop and participants must nominate to the system operator their physical intentions for production and consumption over the subsequent delivery period. In many markets this represents the point ("gate closure") at which voluntary bilateral forward trading between producers and retailers stops and the system operator becomes counter-party to any further real-time trades; in other markets where mandatory engagement in optimized dispatching procedures is required (e.g. markets setting locational marginal prices) this is the point at which final dispatch instructions are received. In either case, thereafter, there will usually be a real-time "participant imbalance", defined

a Corresponding author. London Business School, Regent's Park, London NW1 4SA UK. E-mail: dbunn@london.edu.

b Technical University of Vienna, Austria.

as the difference between a prior nomination (to produce or consume) and the subsequent metered volume, and the aggregate of all of these, the "system imbalance" must be managed through the real-time balancing directives of the system operator. All participant imbalances must be settled financially ex post via administered "imbalance prices".

This terminology for "imbalances" is becoming quite widespread in the industry. The mechanism for setting the imbalance prices is evidently not a market in the sense that the prices emerge endogenously from producers and consumers agreeing to trade between each other, but they are administered following well-defined procedures by the system operator on the imbalance volumes of the participants, and may, depending upon the arrangements, constitute costs or revenues to the participants. Conventional wisdom used to be that, for system control purposes, market participants should be obliged to keep to these real-time nominations, either through central control or motivated to do so by the application of penalties on their imbalance volumes. However, it is an open question if further liberalization, involving a relaxation of this obligation in order to permit or even encourage a degree of participant imbalance would be beneficial, and if so, how might market participants manage their operations accordingly.

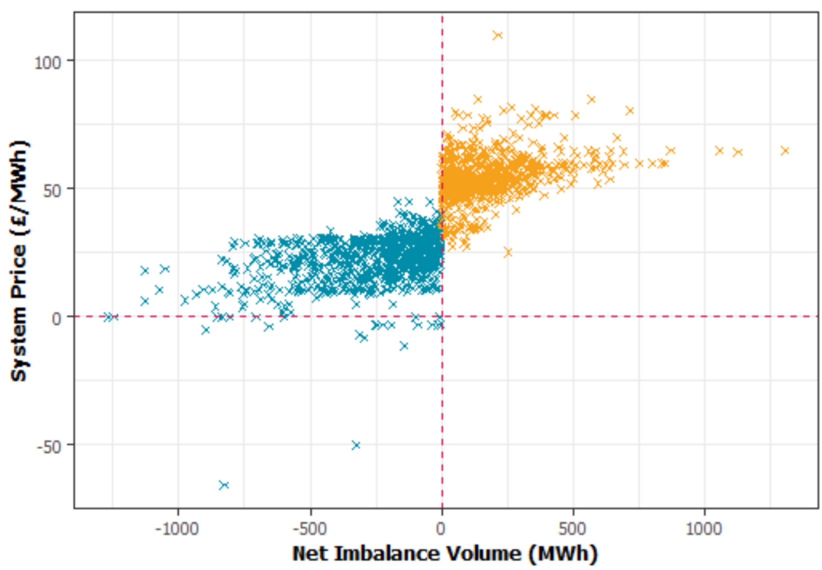
The economic efficiency argument for this relaxation is appealing. If the settlement price faced by the participants for their imbalances reflects the actual-real time cost of the balancing actions taken by the system operator, market participants that correctly anticipate the direction of system imbalance for the whole system can profit by being out of balance in the opposite direction. Thus, if the system as a whole is short (i.e. actual load is higher than expected from the aggregate participant nominations) and the system operator is therefore buying balancing energy, the imbalance settlement price of the system will be higher than the forward market prices, and so a generator that "spills" extra power in the delivery period will receive this higher imbalance settlement price for its imbalance. The situation is similar for active consumers and analogously beneficial in reverse when the system imbalance is long. By attempting to profit from "opportunistic imbalancing," active participants would seek to correctly anticipate the direction of system imbalance and would effectively be undertaking a form of statistical arbitrage (i.e. trading on the basis of statistical expectations). If successful (and permitted), they would thereby gain balancing revenues for themselves and reduce the balancing needs of the system operator.

Such speculation is risky and raises the question of what aspects of market design render the risk-return payoffs for participants taking these actions to be profitable. Beyond that question are the obvious system operations concerns that if there were to be excessively synchronized multi-agent responses to real-time price signals, lagged to some extent by information flows, the stability of the network may be put at risk. In this paper, therefore, we analyse the incentives for both asset-backed physical traders and non-physical ("virtual") traders to engage in this statistical arbitrage through opportunistic imbalancing, how this might become more profitable as short-term wind and solar forecast errors increase the level of system imbalances faced by the system operator, and whether these activities can reduce the overall balancing costs to the system operator whilst limiting any potential detrimental effects on system instability. Note for clarity that this "opportunistic imbalancing" is distinctly not a market power strategy by a large player influencing quantities and prices in the spot market; rather it is a more benign price-taking response by any player, large or small, to an expectation that the system operator for the market as a whole will be buying or selling in real-time. This is not envisaged as a strategic game between a market participant and the system operator.

The first pre-condition to attract this statistical arbitrage is evidently a single price settlement process, ie for each real-time delivery (imbalance settlement) period, the system operator defines one imbalance price that will be applied for settlements in both directions: if a participant is spilling it will receive this price for its imbalance volume whereas if it is in deficit it will pay this price for its imbalance volume. In markets where real-time balancing is undertaken through repeatedly optimized dispatching adjustment algorithms, eg every 5mins as in California, New York and various other ISOs, these final-run, real time prices are naturally used as the basis for settlement (Wang et al., 2015). In other markets, the system operator may be calling upon reserves for up or down regulation minute by minute, and/or accepting bids and offers from generators and consumers continuously during each delivery period, eg in Britain and Germany.

In the continuous case, the system operator may be taking balancing actions in opposite directions many times during each delivery periods. In these circumstances, the duration of imbalance settlement periods may be hourly, half hourly or 15 mins and so identifying a single imbalance price for settlement is usually quite detailed and specific. Usually some procedure is followed to identify the net imbalance volume over the imbalance settlement period and estimate the corresponding marginal cost of the system operator's balancing actions to determine the imbalance price (e.g.. Elexon, 2016). Figure 1 shows the relationship of the imbalance prices in GB in July 2019 (called "system prices" by the System Operator) to the net imbalance volumes in each half-hourly delivery period. Depending upon whether the system is short (positive imbalance) or long (negative imbalance) the prices show distinct distributions. Fundamentally for a short market, the price will be above marginal cost (since the system operator is buying marginal power), whilst for a long market, the price will be below marginal cost (since the system operator is selling power to reduce marginal production).

Figure 1: Imbalance Prices and Imbalance Volumes in GB in July 2019 (Source: Elexon)



In contrast, several markets use dual settlement prices derived from this continuous process to represent the different average, or marginal, costs of buying and selling actions by the system operator during each delivery period (e.g. France, Spain, Italy). Dual settlement prices are generally applied in order to deter the opportunistic imbalancing, as indicated above. In this case, to the extent that the system operator's buying/selling prices will be above/below the market price before delivery, it would always be better for a participant to buy/sell in the forward market pre-delivery, rather than seeking to be imbalanced during the delivery period.

To motivate our analysis of the benefits or otherwise of statistical arbitrageurs operating in the balancing and settlements process, we record a “natual experiment” in the progressively liberalized evolution of balancing arrangements in the British wholesale market. When the British market was first liberalized in the 1990s, central control was retained with a mandatory day-ahead competitive auction providing an algorithmic unit commitment and disptach, followed by socialized balancing costs recovered through an “uplift” in the wholesale purchase costs for the retailers. When the trading was further liberalized to voluntary bilateral transactions in 2001, dual settlement prices were initially implemented to incentivize self-balancing by market participants.

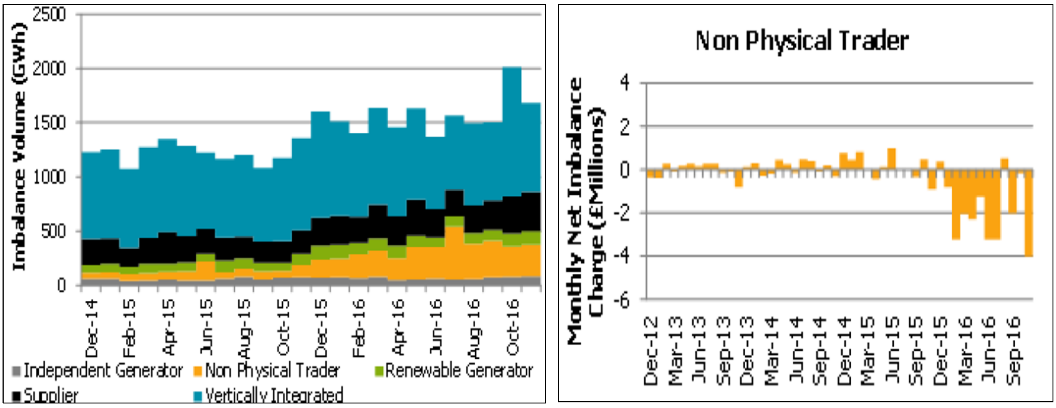
Shortly afterwards, a modification was introduced to avoid penalising those imbalances that were beneficial to the system (i.e. participant imbalances counter to the direction of system imbalance) by having one of the dual settlement prices revert to the pre-delivery, forward market price if it were in the counter-direction to the system. Whilst avoiding a penalty, this still did not incentivise opportunistic imbalancing. Then in 2015, the dual settlement system was changed to a single price, mainly to provide a clearer signal for the provision of flexible reserve capacity and innovative services. In proposing this, the regulatory body noted that the previous dual pricing ‘*drives inefficiency in balancing by over-incentivising parties to balance’ and that ‘under a single price parties with reducing imbalances benefit from the cost saving their reducing imbalances deliver for the System Operator*’ (OFGEM, 2014).

However, despite this rather clear regulatory message, there were mixed signals on “opportunistic imbalancing” as it was still a license condition that large generators should not deliberately take actions to be out of balance. However, this condition was not imposed upon retailers because of the uncertainty in the behaviour of their customers and not thereby imposed upon their distributed energy resources (e.g. small embedded generating units) or upon their demand-side response activities. Hence, there is an opportunity with single imbalance settlement price systems for a vertically integrated company to use real-time demand response contracts with its retail customers to hedge any real-time production problems. For example Dong Energy (now Orsted) initiated a “Renewable Balancing Reserve” contract to commercial customers in which it hedges output deviations during settlement periods from its large wind farms by activating demand response with commercial customers. The company is thereby out of balance on both its retail and generation accounts, but the associated costs and revenues balance out for the integrated company (Brendan, 2016). Improving the risk management flexibility options in this way may indeed be the most persuasive case for liberalising the balancing requirements. Furthermore, to the extent that speculators generally improve market efficiency by providing extra liquidity, “non-physical” market participants (virtual traders) have also been permitted to engage in deliberate out of balance positions. Essentially they would leave open positions from prior trading in the power exchange to be settled against imbalance payments .

Thus, following the change from dual to single imbalance pricing in November 2015, we see in Figure 2 (left) that the system imbalance volumes for the “non-physical traders” in particular started to increase substantially. Furthermore in Figure 2 (right) we see that this became profitable. They had paid a net imbalance charge of £1.94 million in 2014/15, which became a net credit of £21.4million in 2015/16 leading to a profitability of about £10/MWh, or a profit margin of about 20% on the average power exchange prices at the time. More recent analysis (Elexon, 2019a) has shown a steady increase in the ratio of positions taken against the direction of the market by the non-physical (i.e. financial) players.

Motivated by this circumstantial evidence, the research in this paper seeks to analyse the potential effectiveness of this statistical arbitrage more formally. We specify optimal decision-mak-

Figure 2: Imbalance Volumes by Participant and Imbalance Charges for Non-Physical Traders



Source: Elexon, 2017

ing by physical and non-physical participants on the basis of realistic ex ante forecasts. By means of quantile regressions we predict the conditional distribution of the system imbalance and presume that these participants will take optimal expected value positions on deliberate imbalance spillage or shortage. We consider time-delayed information lags in the forecast (15 minutes to 120 minutes) of the system imbalance, as well as wind and solar generations. We chose to calibrate and back-test the model on data from the Austrian zone of the German/Austrian market. The reason for choosing Austria is because there is a well-defined formula for deriving the single imbalance settlement prices and the data is sufficiently transparent. In contrast, an analysis of the British market would involve many behavioural assumptions to model price formation during the continuous bid-offer acceptances in real-time. Furthermore, it was apparent at the time that the statistical arbitrage being envisaged was not against the Austrian (but would have been against the German) market rules.

The paper is structured as followed. In the next section we review briefly the limited research related to this topic, then, in sections 3 and 4 we clarify the market design and the role of the agents in the simulations. In section 5 the model set up is described and in section 6 the results are presented. Overall, the contribution of the paper is in three research directions. Firstly, we demonstrate that operational flexibility, if permitted, can use the predictive value of timely market information to extract value from optimising an imbalanced position in the balancing market. Further, from a market design perspective, we conclude that market liberalisation to permit agent optimisation of these positions appears to be beneficial to the system as well as for the agents, whether physical or speculative. Furthermore, the case for more timely and transparent information on the state of the system is supported by our analysis.

2. BACKGROUND RESEARCH

Imbalance and settlement issues have started to attract an increasing amount of research, mainly because the need for balancing is growing with the substantial introduction of intermittent renewable resources and the growth in demand-side engagement. Related in context to our research, Vandezande et al. (2010) focused on the design of balancing markets in Europe with the penetration of wind power, van der Veen and Hakvoort (2016) identified the relevant variables and performance criteria that play a role in the design and analysis of European balancing markets, van der Veen et al. (2012) analysed the impact of alternative imbalance pricing mechanisms on balancing market performance, and carried out an agent-based simulation using reinforcement learning, whilst Möller et al. (2011) determined strategic positions in the balancing market and identified corresponding economic incentives in Germany.

More generally, the provision of flexibility and demand-side management has been studied extensively as a response to increased balancing requirements, eg reducing reserve requirements through demand side management (Kies et al., 2016), the impact of renewables on flexibility needs (Koltsaklis et al., 2017), the market design requirements (Bertsch et al., 2016), the benefits of solar power forecasting for imbalance markets (Kaur et al., 2016) and optimal bidding of wind production (Pinson et al., 2007). Using hierarchical optimisation, Dowling et al. (2017) show that balancing services can contribute more than half the revenue stream from flexible generating units, and this is a value trend that is increasing. Strategic positioning ahead of the balancing market has also attracted research. Ding et al. (2017) proposed a two-stage stochastic model for an integrated price-maker strategy, where uncertainty of wind power generation and balancing prices are taken into account. Krishnamurthy et al. (2018) formulated energy storage arbitrage under day-ahead and real-time price uncertainty leading to a novel stochastic bidding approach.

Closer to our specific focus, statistical arbitrage between day ahead and balancing markets was analysed in Boogert and Dupont (2005). The authors found that the associated risk of loss for arbitrage potential 24 hours ahead is significant and profits were rarely positive. On the other hand, with shorter lead times, Weber and Just (2015) have indicated that arbitrage between the intra-day and imbalance settlement prices in Germany has been attractive and imbalance positions have been exercised to a cautious extent apparently against the market rules. They observe that has been mainly due to distinctly different price formation processes for the intra-day spot and imbalance markets providing a systematic degree of predictability to the arbitrage. Browell (2018) derives revenue-maximising and risk-constrained strategies for stochastic generators participating in electricity markets with a single price balancing mechanism and Lisi et al. (2018) forecast the direction of system imbalances in the Italian market zones. Nevertheless, a back-tested analysis of the processes that would favour a more liberalised approach to the market arrangements for imbalance, and whether it would be beneficial to both the market participants and the system, has not apparently been thoroughly analysed in research publications.

3. MARKET MECHANISM

The German/Austrian intra-day power exchange is operated by EPEX Spot SE and the power market area comprises 5 delivery zones managed by 5 Transmission System Operators (TSOs), one of which is the Austrian Power Grid (APG). Intraday-trading occurs continuously 7 days a week and the five delivery zones are traded from one order book. The basic intraday delivery period is 15min which can be traded until 30 minutes before delivery begins. For the Austrian delivery zone, internal schedule changes (within the zone) are allowed up to 15mins before delivery (but international flows require 45min notice).

APG is part of the synchronised European grid and follows the standard process of acquiring "control" (i.e. reserve) power to ensure the frequency and operational security. Primary and secondary control power is deployed automatically within 30secs and 5mins respectively, whilst tertiary is activated by the TSO on a 15min basis to replace any substantial use of secondary reserves. Tertiary control is used to relieve secondary control so that this is once again free to support or restore the availability of primary control, should this be necessary. The price of control power is set by weekly and daily auctions. In the daily auctions, an energy price for tertiary control, p^{tert}, is set on a pay-as-bid basis.

The system imbalance *imb* for a particular settlement period *j* is the amount of energy (positive or negative) required by the system operator to keep the demand and supply in balance, and thus the system frequency at 50Hz. Any mismatch, ie system imbalance, suggests that demand and supply are different than expected. Thus, the system imbalance (*imb*) for a settlement period is a deficit of production, ie "short" if *imb* is +ve, or a surplus of production, ie "long" if *imb* is -ve. A deficit, positive *imb,* is met by the system operator calling on extra reserve production; whereas a surplus, negative *imb,* is sold back to some flexible market participants, presumably at a price that is lower than their marginal costs and thereby causes them to reduce production. APG publishes every 15mins the preliminary estimate of system imbalance, with a lag of 10mins. The single imbalance price for settlements, $p|imb$, is determined in Austria by means of a formula which consists of a "basis price" and a "transfer function". The basis price p^{Basis} can be positive or negative, defined as:

$$p^{Basis}\begin{cases}\min(p^{tert},p^{ID},p^{DA}) \text{ for } imb<0 \text{ and activated tertiary}\\ \min(p^{ID},p^{DA}) \text{ for } imb<0 \text{ and no tertiary}\\ \max(p^{tert},p^{ID},p^{DA}) \text{ for } imb>0 \text{ and activated tertiary}\\ \max(p^{ID},p^{DA}) \text{ for } imb>0 \text{ and no tertiary}\end{cases} \tag{1}$$

where p^{ID} is the hourly average intraday price for that 15min period as traded previously on the wholesale power exchange (EPEX Spot), $p^{DA}$is the previous relevant hourly day ahead auction price (administered by EXAA) and p^{tert} is the volume-weighted average price for any activated tertiary control power in that 15min delivery period. The nonlinear "transfer function" is defined as

$$T=\min\left(U_{min}+\frac{U_{max}-U_{min}}{imb_{max}^2}\times imb^2;U_{max}\right) \tag{2}$$

where, in our case-study, $U_{max}=40\,EUR/MWh$ and $U_{min}=3\,EUR/MWh$. These are the constant maximum and minimum parameter values of the transfer function *T* for the monthly data in our analysis.

The ex post calculated imbalance price is then:

$$p\,|\,imb=p^{Basis}\ \pm\ T \tag{3}$$

Depending on the state of the imbalance the imbalance price function follows a quadratic term within the range $\pm imb_{max}$ and is constant for the outer ranges $|imb|\geq imb_{max}=70\,MWh$. The transfer function is positive if the imbalance of the system is positive (i.e. the system imbalance is short) and vice versa negative. This is illustrated in Figure 3.

Table 1 shows the system imbalance data for February (winter) and August (summer) in 2015. The system imbalance was, in the summer month, on average 3 MWh short. In that month the system was short in 1758 periods with a cumulated imbalance of 35 GWh and long for 1218 periods with a cumulated system imbalance of 26 GWh. In contrast, in the winter month, the system imbalance was on average 7 MWh long. In that month the system was short in 1097 periods with a cumulated system imbalance of 20 GWh and long for 1218 periods long with a cumulated control power requirement of 41 GWh.

Figure 3. Balancing price formation (source: APCS).

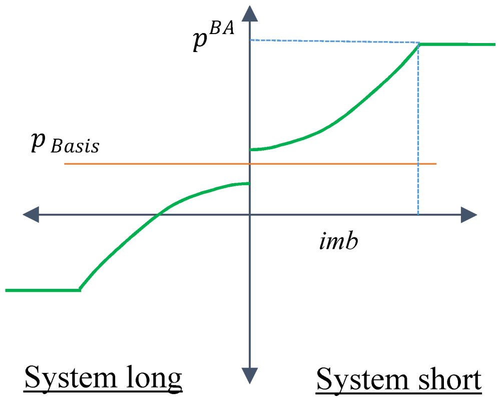


Table 1: Imbalance data February (winter) and August (summer) 2015

summer	imbalance [MWh]	imbalance price [EUR/MWh]	winter	imbalance [MWh]	imbalance price [EUR/MWh]
max short system	123.95	247.66	max short system	84.19	238.47
min long system	–160.24	–475.36	min long system	–172.29	–327.62
arithmetic mean	3.05	30.84	arithmetic mean	–7.79	24.51
standard deviation	27.71	47.11	standard deviation	29.89	70.47

As in Eq (1), the imbalance settlement prices are sensitive to the activation of tertiary control power, but, in that month, tertiary was activated quite infrequently. Prices for the activation of tertiary were within a range of –400 EUR to + 250 EUR depending on the magnitude of activated tertiary control power.

4. STATISTICAL ARBITRAGE MODEL

In the above section, a positive imbalance at the system level means that the system is short of generation. This will be the net effect of all the individual participants' imbalances. Thus, if an individual generator's imbalance, x, is positive, it is short (i.e. it nominated more than it generated), its imbalance charge will be $(p|imb)x$. If the system as a whole is short, $p|imb$ will be positive and the generator will be paying the charge, $(p|imb)x$ as a cash outflow. But if x is positive and the system is long, $p|imb$ may be negative and its imbalance charge $(p|imb)x$ would be negative. That will represent a cash inflow for the generator. Likewise, the imbalance charge will be negative (cash inflow) for negative x and a positive $p|imb$, when the generator is spilling power (i.e. generating more than it nominated) and the system as a whole is short. Recall Figure 2 where the charges in Britain switched from positive to negative as traders began to profit from speculating against the direction of system imbalance.

Thus, if a market participant has a rational expectation $\widehat{imb}$ for the system imbalance (based on the latest information on the system imbalance from the TSO and expected values for solar and wind power forecast errors) and deliberately intends to take an imbalance position defined as x, it would be able to anticipate a conditional imbalance price expectation $\hat{p}\,|\,\mathrm{x},\widehat{imb}$. For a particular settlement period $j \in \{1, \ldots, T\}$, the system imbalance $\widehat{imb}$ can be forecast and thereby modified by a market player´s action x. An adjusted system imbalance $\widehat{imb} + x$ implies a change from $\hat{p}\,|\,\widehat{imb}$ to an imbalance price, $\hat{p}\,|\,\mathrm{x},\widehat{imb}$. Including x in (2) leads to the imbalance settlement price:

$$\hat{T}(x) = \min(U_{min} + \frac{U_{max} - U_{min}}{imb_{max}^2} * (\widehat{imb} + x)^2 ;\ U_{max})$$
$$\hat{p} \mid \mathrm{x}, \widehat{imb} = \hat{p}^{Basis} \pm \hat{T}(x) \tag{4}$$

Referring to (1), and for a particular imbalance settlement period j, we estimate $\hat{p}_j^{\text{Basis}}$ from the EPEX spot reported average intraday price $\hat{p}_j^{\text{ID}}$ shortly before gate closure and the day ahead price p_j^{DA} from day-ahead auctions. Since tertiary is hard for market participants to predict and was activated only 48 times in our dataset, we assume pragmatically that agents may generally not seek to anticipate its effect in this conditional price expectation (we backtest under this assumption later).

The market participant´s behaviour is influenced by the expected spread between the participant's marginal cost *mc* (which may be the forward market price for a trader, or the short-run marginal cost for a generator) and the expected imbalance settlement price. The payoff value (cash inflow), *v*, to the participant of taking an imbalance position *x* is:

$$v = (mc - \hat{p} \mid x, \widehat{imb})x \tag{5}$$

Thus, positive *v* (cash inflow) can occur with either positive or negative positions, *x*. If *mc* is positive and higher than the expected imbalance settlement price (which may be negative), the player would want to be short with positive *x*. Alternatively, if $(mc - \hat{p} \mid x, \widehat{imb})$ is negative and the player goes long (over produces compared to nomination), *x* is negative and again *v* becomes positive (cash inflow).

Regarding the marginal cost in (5) we consider two different market players:

Physical player

Consider a part-loaded thermal player who has nominated a production schedule before gate closure and who is able to adapt production output (up-regulation and down-regulation) with short-run marginal cost, *mc*, in €/MWh electricity. We presume the generator to be operating, either because it is in merit, or because it is in a must-run regime through local system conditions or CHP obligations to an industrial or metropolitan site. We consider a simple characteristic model of a gas turbine with efficiency $\eta = 0.5$, and market prices for the gas. In Austria, a two price system for balancing gas is in place and the gas imbalance settlement costs have a mark-up of plus or minus 3% on day-ahead gas prices ($p^{gas_{long}} = 1.03 p^{gas_{DA}}; p^{gas_{short}} = 0.97 p^{gas_{DA}}$), or in case of higher imbalances, on a volume weighted mean value for gas balancing costs. Variable maintenance costs are not considered in this model, whilst electrical transmission and related electrical fees are small and substantially paid by consuming rather than producing units.

For overproduction (participant going "long") compared to nomination, the payoff value is determined by the price difference of the expected imbalance settlement price $\hat{p} \mid x, \widehat{imb}$ and the marginal cost, $mc^{gas_{long}}$, where

$$mc^{gas_{long}} = \frac{p^{gas_{long}} + p^{CO2} + p^{grid} + p^{taxes}}{\eta} \tag{6}$$

where $p^{gas_{long}}$ is the price of balancing gas, p^{CO2} is the cost of the necessary carbon allowances in the EU ETS, p^{grid} is the use of gas grid system charge and p^{taxes} are the various levies on the natural gas usage. If $\hat{p} \mid x, \widehat{imb} \rangle mc^{gas_{long}}$ the physical player is incentivized to spill *x* with payoff:

$$\mathrm{v}^{long} = (mc^{gas_{long}} - p \mid x, \widehat{imb})x \tag{7}$$

For considering the short position in the balancing market, it is only relevant to consider the incremental value of selling balancing gas for the avoided production:

$$mc^{gas_{short}} = \frac{p^{gas_{short}} + p^{CO2} + p^{grid} + p^{taxes}}{\eta} \tag{8}$$

Recall that a two price system for balancing gas is in place and the gas imbalance settlement costs have a mark-up of +–3% on day-ahead gas prices or in case of higher imbalances, on a volume weighted mean value for gas balancing costs. A short position therefore has marginal payoff, with the day ahead trading margin being sunk:

$$\mathrm{v}^{short} = (\, mc^{gas_{short}} - \;\hat{p} \mid x, \widehat{imb})\, x \tag{9}$$

The TSO publishes the latest information on system imbalance 10 minutes after the previous imbalance settlement period. Based on that information and the latest wind and solar generation forecast errors, the physical player can make a decision to up- or down regulate the thermal power plant. We therefore assume a decision lead-time for a physical player to be 30 minutes, ie a rational expectation is lagging by two periods (including the delivery period itself).

Non-physical player

A non-physical player is a trading company who is active on the wholesale power exchange (EPEX Spot) but does not physically control production or consumption after gate closure. A non-physical trader contracts with physical participants and thereby influences their generation or consumption. This agent buys or sells on the power exchange of the Austrian/German market zone, but does not close out its position with a counterparty (retailer or generator), rather it is left open, exposed to an imbalance charge, in order to speculate upon a margin from the imbalance price in the settlement process. The pay-off function of the non-physical player is therefore the same for both short and long positions:

$$v = \left(mc^{EPEX} - \hat{p} | x, \widehat{imb} \right) x \tag{10}$$

We assume that the non-physical player is able to buy/sell last price EPEX spot (at mc^{EPEX}) in the Austrian/German intraday market with delivery in the Austrian TSO zone. This is facilitated by the high market liquidity due to the integrated German/Austrian intraday market comprising 5 delivery zones. Table 2 provides price data for the two backtesting months in our study (February and August). Evidently the average price spreads between the power exchange and imbalance settlement prices provides intuition for the potential arbitrage by speculators. Thus, the lower average imbalance settlement prices compared to the final exchanged traded prices may also tempt more short compared to long speculative position. The volatility, and hence risk, is high, however.

With regard to the time-line for the non-physical player, the crucial difference with the physical player is that the non-physical player is constrained by the Austrian regulations for nomination of cross border schedules which requires a gate closure at 45 minutes before delivery and a delivery time of a full hour (APG (2016)). The variable x therefore needs to be taken for 4 consecutive 15min imbalance settlement periods, and our time steps are hourly for the nonphysical trader. The TSOs processing lag to publish the latest information on the system imbalance is 10 minutes

Table 2: Average Prices and Volatility in the two backtested Summer and Winter months

Monthly prices in [EUR/MWh]	Day Ahead Price	Volume Weighted Price EPEX intraday	Last price EPEX intraday	Imbalance settlement price
average price (summer)	31.55	34.46	35.77	30.84
standard deviation (summer)	8.83	14.32	25.10	47.11
average price (winter)	32.50	36.09	36.43	24.51
standard deviation (winter)	15.43	14.64	18.06	70.47

as it was for the physical player. This leads to a decision lead time of between 75 minutes and 120 minutes. Evidently these administrative restrictions create a more challenging forecasting problem for the non-physical player.

5. OPTIMAL POSITIONS

We assume risk neutral players seeking to maximise expected value. Let $\left\{\widehat{imb_i}\right\}$ represent a set of expected system imbalance states (in MWh), being a discretisation of the system imbalance variable, such that $prob(\widehat{imb_i})$ is the probability of occurrence of each estimated state and x_k are the discrete actions (deliberate spillage/shortage decisions) in MWh with index k that can be taken by the market player. Then for every imbalance settlement period $j \in \{1, ..., T\}$ a spillage or shortage decision $x_{j,k}$ which maximizes expected outcomes, is chosen. The time steps are 15mins and we consider a decision-making lead time of up to two hours ($T = 8$). The decision $x_{j,k}$ is dependent on the system imbalance estimates for every imbalance settlement period j, $\widehat{imb_{i,j}}$, and the anticipated price response to $x_{j,k}$, $\hat{p}(x,\widehat{imb})_{i,j,k}$. For each state and action the corresponding payoff value is therefore:

$$v_{i,j,k} = (mc_{i,j,k} - \hat{p}(x,\widehat{imb})_{i,j,k})x_{j,k} \tag{11}$$

Hence the optimal expected value action is:

$$x_j^* = \max_k \left(\sum_{i \in I} v_{i,j,k} * prob(\widehat{imb_i}) \right)_j \tag{12}$$

In order to model the conditional dependence of the system imbalance probability distribution on the short term flow of information to the agents, quantile regression (QR) was used (details in the Appendix), specified as:

$$\widehat{imb_{i,j}} = \beta_{i1,j-r} * imb_{j-r} + \beta_{i2,j-r} \times wind_{j-r} + \beta_{i3,j-r} * solar_{j-r} \times D_s + c_{i,j-r} \tag{13}$$

where:

imb_{j-r} is the system imbalance variable with a time lag of r,

$wind_{j-r}$ is the wind forecast error, calculated as the difference between the day ahead forecast and the latest value measured at (j-r), and

$solar_{j-r}$ is the solar forecast error, calculated as the difference between the day ahead forecast and the latest value measured at (j-r). Since the solar error was significant only in summer, there is a binary variable D_s which is 0 in winter and 1 in summer.

To undertake a back-testing analysis and evaluate the system behaviour based on measured system imbalance data, following the estimations based on January and July, the months of February and August 2015 were evaluated as out-of-sample simulations of the statistical arbitrage trading.

For the evaluation of the deliberate spillage/shortage positions, we compared evolving system costs, profit potential and measures for the volatility of system imbalances. As baseline, in the absence of statistical arbitrage actions, we defined the system costs for balancing as:

$$system\ costs^{base} = \sum_{j \in T} p_j \mid imb_j^{real} * imb_j^{real} \tag{14}$$

With statistical arbitrage, the balancing system costs in the optimized case ($system\,costs^{opt}$) are determined by the imbalance settlement price function $p_j \mid x_j^*, imb_j^{real}$ adjusted by the optimized system imbalance due to long/shortage position $imb(x)_j^{opt} = imb_j^{real} + x_j^*$.

$$system\ costs^{opt} = \sum_{j \in T} p_j \mid x_j^*, imb(x)_j^{real} * imb(x)_j^{opt} \tag{15}$$

The profit, ie payoff value, for the market participant is defined as

$$profit = \sum_{j \in T} (mc_j - p_j \mid x_j^*, imb(x)_j^{real}) x_j^* \tag{16}$$

In consideration of the general size of system imbalances, as shown in Table 1, we set the maximum long position for x_j^* at 60 MWh and the maximum short x_j^* at 100 MWh.

The effect of volatility and the impact of time delayed decisions are measured by the monthly cumulated absolute system imbalance and the system imbalance standard deviation before and after optimization. Furthermore, to investigate potential dynamic instability due to lagged responses, following Larsen et al. (2015) we also use system imbalance half cycles, measured as switches between positive and negative system imbalances, as an indicator of system imbalance cycling.

Since the physical and the non-physical agents face different time lags, we undertake a dynamic analysis for both players. For that purpose we vary information time delays for the decision model starting with lag 1 up to lag 8 forecasts and run simulations for summer and winter to study the effect of time delayed decision making. Even though the very short lags may be infeasible under the current Austrian market rules, we considered the full range of lags in order to inform a debate upon the benefit of market rule changes and greater transparency.

6. RESULTS

This section consists of three parts. First the results of the forecasting models are presented. Then the physical and non-physical player's profitability based on optimal decisions from the forecasts are compared to a perfect foresight simulation. Thirdly the impacts of time delayed decisions in a dynamic setting are evaluated.

Forecasting model

The estimation results can be summarized as follows:

$\mathrm{imb}_{i,j-r}$ had significant positive coefficients for the whole range of quantile regressions and time delays. As expected, this reveals the adaptive market conditions at 15min imbalance settlement periods.

$wind\_e_{i,j-r}$ had negative coefficients, as expected, being significant for almost all quantiles. In summer and winter for quantiles above q0.6 wind production error has no significant impact if the time delay is longer than 5 intervals. This implies that there may be dynamic corrections at these lags.

$solar\_e_{i,j-r}$ had significant negative coefficients in summer except for those quantiles above q0.7. Those higher system imbalances are presumably caused by other factors.

Figure 5: Comparison of Lag 1 and Lag 8 Prediction Intervals

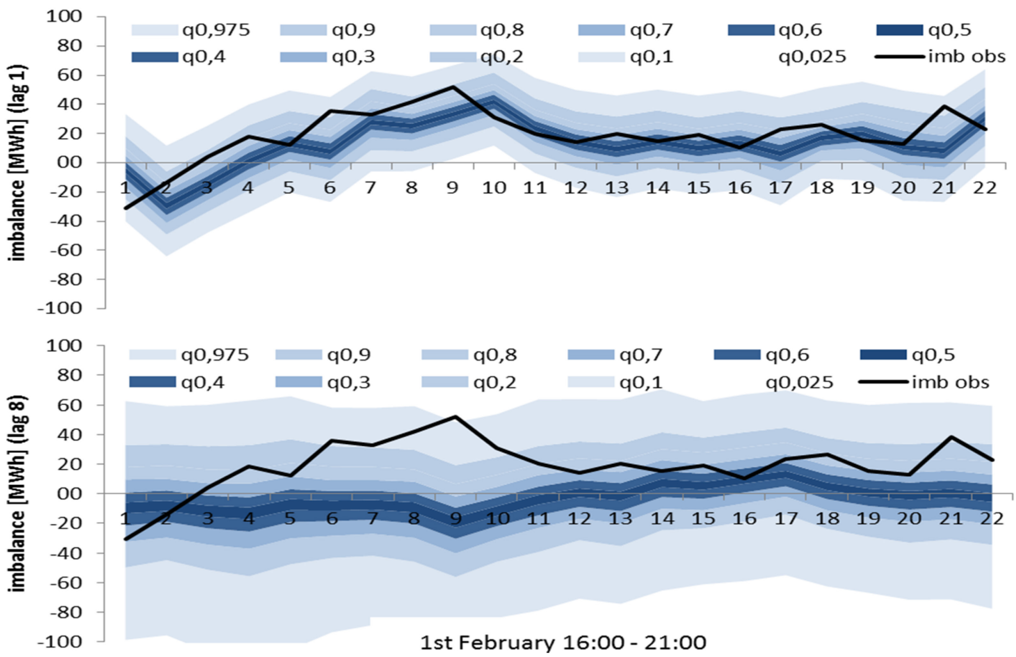


Figure 5 shows the prediction intervals from the lag 1 and the lag 8 quantile regression models on 1st February 2015 between 16:00 and 21:00, compared to the observed system imbalance shown in black. The average range of the quantiles (q0.975 and q0.025) for the lag 1 (15 mins information time delay) is 45 MWh. As expected, at 8 lag (120 mins information time delay) the model adapts slowly and the range of the quantiles (q0.975 and q0.025) is on average 103 MWh, but still apparently accurate in coverage. The conditional distribution gets wider for the extreme quantiles with the longer time delay.

Back-test Simulations

Backtesting was based upon summer (August) and winter (February) months, simulated for each 15min delivery period. The forecasts are based upon models estimated over the previous months, July and January respectively. In Figure 8, we see how much opportunistic trading was undertaken by the physical and non-physical players in our winter simulations. Evidently the non-physical player is much more active. The non-physical trader buys/sells power at last prices in the EPEX spot market. The higher volumes of trades are caused by the fact that every small price difference between intraday market and expected price from the imbalance settlement is arbitraged by the non-physical player. On the other hand the physical player (e.g. gas CHP unit in part load due to a must run criteria) needs higher incentives / price signals, due to the gas and CO2 prices. Furthermore, the gas turbine generation needs to be in merit which is not all the time.

Figure 8: Cumulative short and long positions in winter 2015

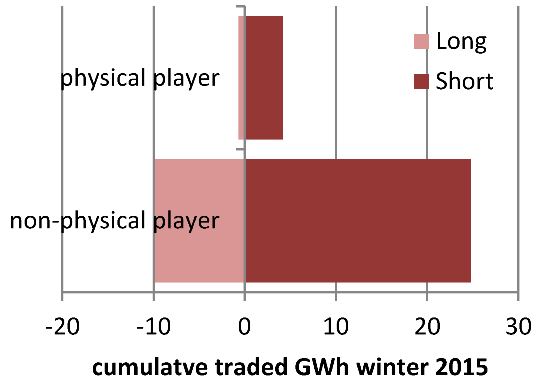


Table 3 gives a summary of the physical and non-physical participants' profits and their effects on system balancing costs for backtesting with quantile regression forecasts (QR) and, for comparison, with perfect foresight (PF).

Table 3: Agent Profits and System Costs using Perfect Foresight (PF) and Quantile Regression (QR)

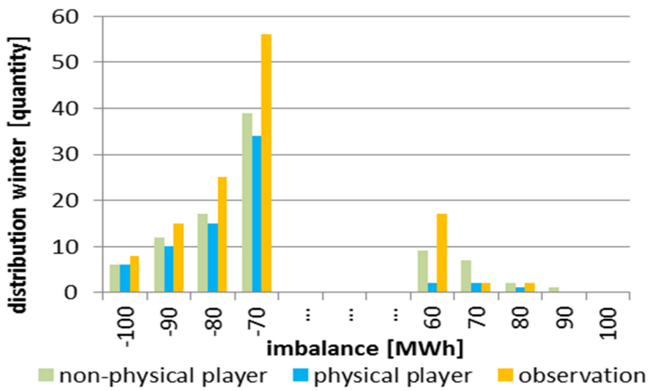
		SUMMER (August)			
	Observed	**QR phys**	**PF phys**	**QR non-phys**	**PF non-phys**
Agent Profits (€)		227,366	330,960	223,045	387,146
System Costs (€)	2,305,338	1,996,222	1,813,617	2,835,883	2,078,886
Standard Dev System Imbalance	27.17	24.96	23.08	29.84	26.88
Absolute System Imbalance	56,267	52,851	48,413	64,083	57,133
		WINTER (February)			
	Observed	**QR phys**	**PF phys**	**QR non-phys**	**PF non-phys**
Agent Profits (€)		503,581	671,902	838,268	2,046,731
System Costs (€)	2,591,436	1,995,951	1,768,552	2,559,272	1,042,558
Standard Dev System Imbalance	29.88	27.19	25.35	31.06	19.60
Absolute System Imbalance	62,848	58,182	53,429	65,143	32,892

The standout results are that both the physical and non-physical agents make profits through opportunistic imbalancing. In summer, the physical agents using QR forecast make about 70% of what they could with perfect foresight, whereas in winter, they achieve about 75%. For the non-physical trader, similar results emerged in the summer, but in the winter, when there were more spikes and more activations of tertiary, perfect foresight gave very high profits to the non-physical trader. System imbalance costs are reduced in all cases, except in the summer for the non-physical agents using the QR forecasts.

In Figure 9 the upper and lower tails of the system imbalance distributions for winter are plotted for the agents using QR forecasts. We observe that negative system imbalances are reduced by additional shortage positions from both the physical and the non-physical players. For positive system imbalances we see that the non-physical player increased the system imbalance somewhat in the upper tail.

Furthermore the relative profits expressed as a ratio of average MWh traded were substantial, being in winter €89/MWh for the physical player and €24/MWh for the non-physical player, whereas in summer the physical player achieved €52/MWh profitability whilst the non-physical player achieved €7/MWh. Recall from Figure 8 that the nonphysical player is trading a lot more. Also recall from Table 2 that the average traded closing price in the power exchange prior to balancing was €36/MWh in winter. So the nonphysical player is making an average profit in winter of 67% on the spread between power exchange and imbalance settlement prices.

Figure 9: System imbalance distribution winter below –70 MWh and above 60 MWh



The fundamentals for the physical player are different, relying upon the spread between marginal production costs (gas) and the imbalance settlement price. Evidently, optimising this gave even higher profitability (per MWh traded) to the physical player, and of course, in practice, the physical player could also engage in the same speculation as the nonphysical. Whilst the actions of the physical player were evidently beneficial in reducing total system costs and therefore welfare enhancing from both producer and consumer perspectives, the nonphysical player's effects were more detrimental to the system. We note that this could be due to the extra delay in information processing with the physical player acting within 15 mins and the nonphysical acting 60mins ahead. Below, we consider the beneficial effects if the administrative restrictions on the nonphysical player were relaxed so that it could act on information as quickly as the physical player.

Figure 8 and Table 4 show a tendency towards shortage applying to both physical and nonphysical players, and this is despite the underlying market circumstances in which, as we noted earlier, there were 50% more settlement periods when the market was short compared to being long (and to be profitable, a player has to be out of balance in the opposite direction to the market). There is evidently a diffrence in the profitability function and perhaps also in the accuracy of the forecasts. The market fundamentals (average prices in Table 2) indicate the incentive to take higher volumes of short positions than long positions for the non-physical player. The physical player also has more incentive to take a short position. Once the power plant is activated in part load on the day-ahead market it is more profitable to curtail than spill. 86%/72% of the trading volume are shortage deals for the physical/non-physical players in winter 2015. The average size of the positions was about 7 MWh and we checked all positions for being realistic, given the size of this market.

We also observe in Table 4, the increased activity of the nonphysical player. In summer the non-physical player took 2516 positions whereas the physical player took only 324 (out of 2684 possible intervals). Figure 8 showed that effect by cumulative traded volume. Evidently, the incentives for the physical player present fewer opportunities but the payoffs are greater, hence the greater profitability per MWh traded.

Table 4: Short/long positions in winter and summer for the QR

		physical player	non-physical player
short positions	winter	380	1641
	summer	271	1466
long positions	winter	63	867
	summer	53	1050

Information Flow

As an opportunistic process, the flow of information is a key aspect of the microstructure and trading performance. One concern that has attracted research is the effect of information time delays on the stability of the system. For example, (Nutaro and Protopopescu 2009) studied the effect of market clearing time and price signal delays in electric power markets and found that there is an upper limit on the market clearing time and the delay of the price signal beyond which the system becomes unstable. Similarly, (Larsen et al. 2015) argued that activating demand response in existing electricity markets can be unstable, resulting in oscillations and increased volatility in supply and demand due to the cobweb effect. In the context of our research, therefore, to the extent that there may be benefits in further liberalisation to encourage opportunistic imbalancing, a market design question is whether the flow of information needs to be accelerated. Thus, Figure 10 shows the overall system balancing costs for both players in both months and for all time lags simulated. The actual observed system imbalance costs are shown for comparison. Thus, for the physical player simulating 1 lag, system costs were cut by 23% in winter and 13% in summer. In the current context, this means that the physical player would have five minutes to act, unless the information were provided with a lag of less than ten minutes. Figure 10 shows the cumulative absolute system imbalance and standard deviation of the system imbalances. In the cumulative absolute system imbalance, we see similar indications as with the imbalance costs, whist the standard deviations also show a further benefit in terms of reducing volatility. More surprisingly, we see that the conjecture, advanced in the previous section, that the detrimental system performance of the nonphysical player was due to its longer information lag (60mins) does indeed appear to be upheld and furthermore that if the lags can be shortened, it can in fact provide even more system benefits than the physical player. A 36% reduc-

Figure 10: Lagged system costs and profit per traded volume

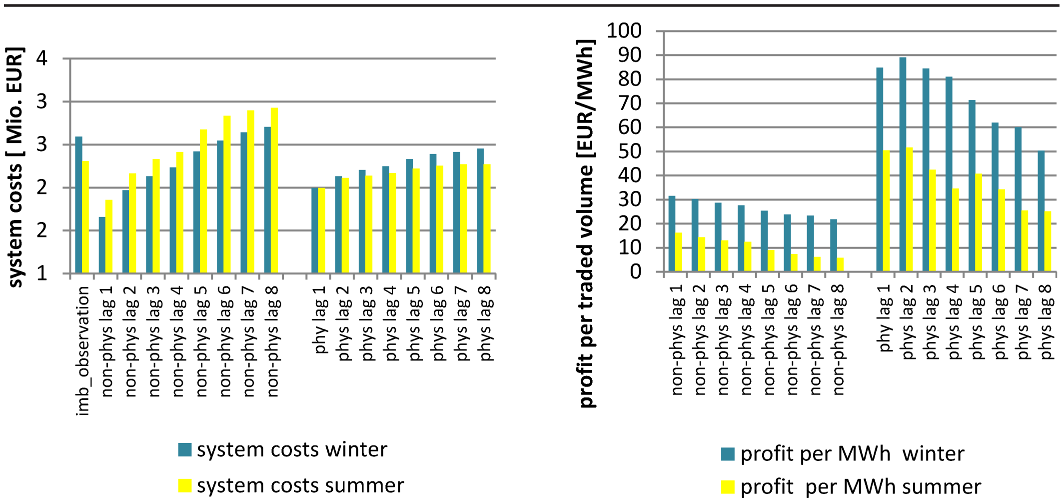


Figure 11: Lagged cumulative absolute system imbalance and system imbalance standard deviation

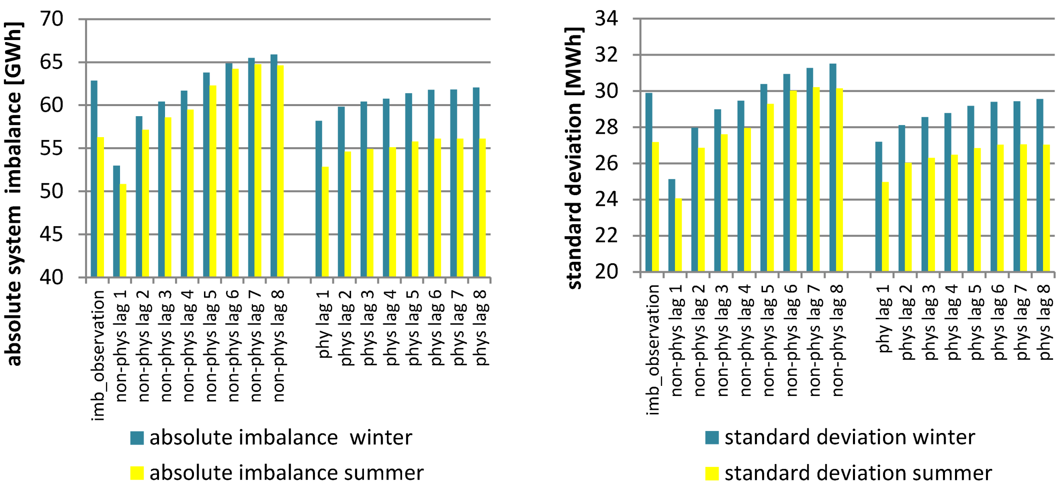
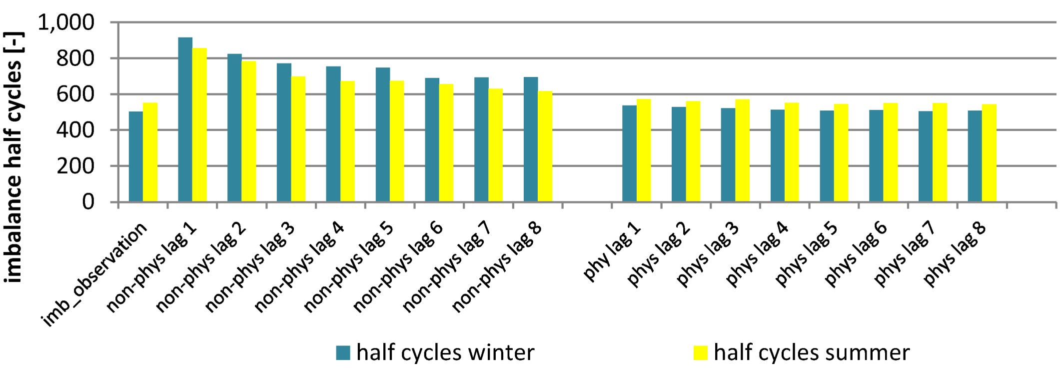


Figure 12: Lagged Imbalance Half Cycles



tion in system costs in winter and 19% in summer can be achieved at lag 1. In contrast, with longer information lags, the effect was more sensitive for the non-physical player and the deterioration of performance more marked at longer lags. Whilst the physical player improved system balancing costs for all lags considered up to 8, the nonphysical started to increase system costs after lag 6 in winter and lag 3 in summer. Profits however, remain positive at all lags considered.

Figure 12 shows the number of system imbalance half cycles (i.e. switches in sign in the time series of 15min system imbalances) for summer and winter. System imbalance half cycles increased slightly for the physical player, but for the non-physical player, the changes of sign of the system imbalance increased substantially. For the lag 1 model we noted an increase of 50–80%, which decreases for longer time delays in winter and summer, but remained above the actual observed data. Expressed in minutes, changes between positive and negative system imbalances were on average 80/73 minutes for the physical player's actions and 44/47 for nonphysical. For more detail, in Figure 12 we compare 100 imbalance settlement periods for the non-physical lag 1 model and the observed system imbalance in summer. It can be seen, that the higher frequency of half cycles are often caused by small overreactions which lead to sign switches.

Figure 13: Summer system imbalance observations and non-physical lag 1 model

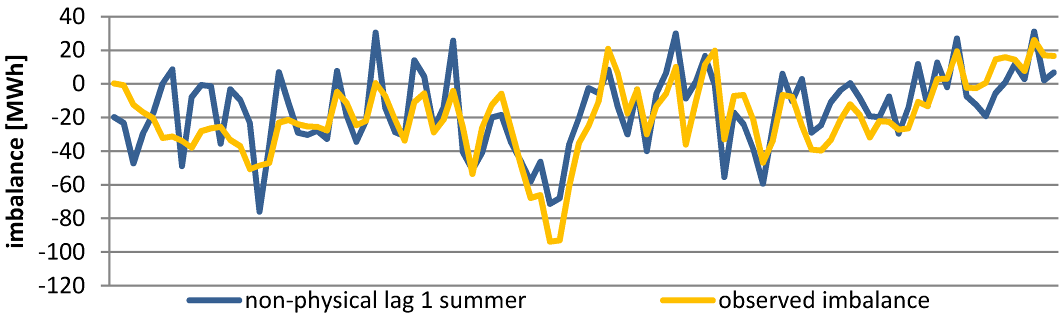
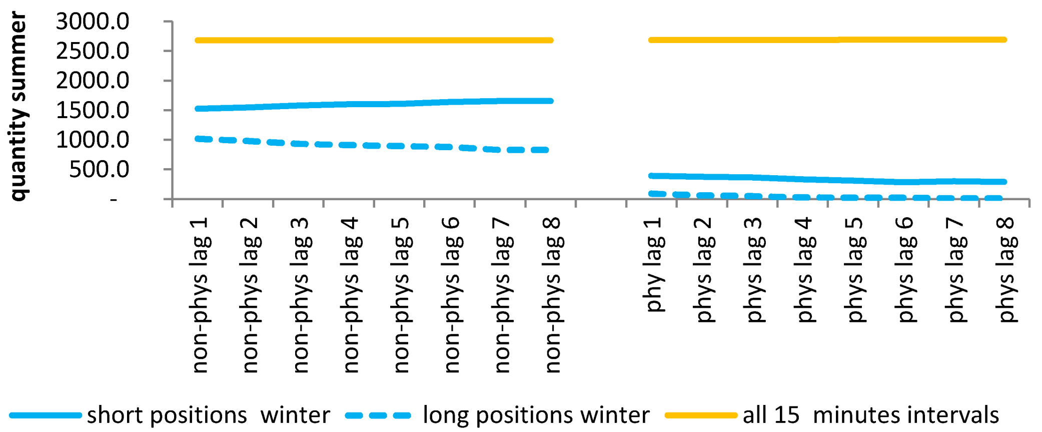


Figure 14: Quantity of short and long positions in winter



It is not obvious whether an increase in system imbalance half cycles is a good or bad effect. Whilst an increase of half cycles could be interpreted as an undesirable indicator from the TSO perspective requiring more activation of control power, one could also take the view that if the system were well balanced in expectation, system imbalance volumes would be a zero mean, stationary process with a high speed of mean reversion, implying a high rate of sign switching. In any case, this increased effect here does not seem to be seriously detrimental.

Finally in Figure 14 we show the quantity of actions in winter 2015. The non-physical player took, in the lag 1 simulation, more than 2500 position (out of 2684 possible imbalance settlement periods). With longer time delays the non-physical player took more short positions (1657 at lag 8), which is rather surprising. In slight contrast, long positions for the non-physical player were taken 1000 times in the lag 1 decreasing to 830 times in the lag 8 model. On the other hand the physical player was more selective. In the previous section, we considered whether the difference in behaviours between the players was due to the nonphysical player having to decide on actions sooner than the physical. Here, controlling for information flow, there is still a substantial difference in behaviour. Thus we have clarified that the difference in behaviour is mostly due to the fact that the physical and nonphysical players face different arbitrage spreads, between imbalance and marginal cost for the physical and between imbalance and closing power exchange price for the non-physical.

7. CONCLUSIONS

Motivated by a natural experiment in the British market design involving a switch from dual to single imbalance settlement prices, we considered the incentive for physical and speculative participants to optimise their expected imbalance positions. Not every jurisdiction permits participants to be deliberately out of balance and so we analysed the behaviour, not only in terms of its profitability for the participants, but in terms of system effects. We conclude that market liberalisation to permit agent optimisation of imbalance positions with timely information below 30 minutes appears to be beneficial to the system as well as for the agents, whether physical or speculative. Furthermore, the case for more timely and transparent information on the state of the system is supported by our analysis. In Britain for example, it is now possible to trade after gate closure up until the beginning of each settlement period, and this final hour has become the most active of all intraday trading (Elexon, 2019b).

For reasons of computational precision, we analysed the effect of statistical arbitrage from a market agent´s perspective (a physical and non-physical player) in the current Austrian policy framework (in which there is a precise formula for the imbalance settlement price) and in a dynamic setting with information time delays between 15 minutes and 120 minutes. For the physical player we found that statistical arbitrage is profitable and has a convincing potential to reduce system costs, absolute system imbalances and imbalance extremes. With long time delays this effect diminishes as forecasting and actions become more difficult. The frequency of trades and imbalance half cycles of the system were observed to adequately stable. We concluded that profitable operations and market efficiency increase by these actions. For the non-physical (speculative) actions, whilst profitable for the player, the system benefits were less substantial and declined with longer information lags. The dynamic analysis shows that up to a time lag of 45 minutes in winter and 30 minutes in summer the non-physical player decreases system costs, system imbalance deviation and absolute system imbalance. The market design implication is that consideration should be given to reducing the extra information lags required for non-physical compared to physical players, as this would be highly beneficial. With a well-designed imbalance settlement price settlement process and timely information flows, agents can thus be incentivised to contribute to stabilizing the power-system, and speculation on the imbalance market should not be discouraged. Finally we should emphasize that in referring to physical and non-physical players, these are simple examples of engagement actions; a company could be involved in both or indeed in more complex portfolio positions (e.g. by also offering tertiary control power to the market).

The main contribution therefore concerns the optimal real-time operations of market participants and the principle of more liberalized engagement. Evidently, as in any arbitrage opportunity which markets present, the more agents that engage, the more the marginal arbitrage benefits become eroded. We have not addressed these multi-agent aspects. In many respects, this aspect is common to other instruments of active real-time responses to market signals, in particular those of demand management, storage or electric vehicle charging. Thus, we also recognize the system stability issue posed by increased real-time voluntary actions in the market, but consider it not specific to this type of activity but rather part of the much broader topic of smart, real-time consumer and producer engagement, which is outside the scope of this analysis. As for the equilibrium implications if a large number of arbitrageurs became similarly attracted to these activities, we have not undertaken a formal analysis. Evidently with risk neutral, rational expectations and arbitrage free conditions, mixed strategies or a focal point of moderate profit sharing may emerge analytically. Evidently a no-trade result would allow the arbitrage to re-appear, and the more aggressive all-trade result to the same market signals would provide negative payoffs for all as the system would become out-of-balance in the opposite direction. In practice, it is more likely that markets will continue to see, as in the UK, Belgium and elsewhere, heterogeneous agents acting cautiously under trading risk limits such that the dynamic interactions will depend upon the transparency of their actions. Nevertheless, this is an open topic for more formal analysis.

Overall, it may seem counter-intuitive that the system may be better served by allowing agents to forecast the system imbalance and trade against its direction, than requiring agents to maintain their own balanced positions and leave the System Operator responsible for centralised balancing actions. The intuition behind our results, which suggest this is not the case, is that System Operator in most jurisdictions activates reserve for balancing in response to the actual state of the system, whereas in our modelling the market agents are more proactive in forecasting the system imbalance states, and thereby take actions which are less costly. Furthermore, the license condition for an independent system operator may preclude them from acting speculatively in the market in the same way as market participants.

There are several technical limitations to this study, but we do not consider them to affect the broad implications. Evidently the agents could be using better predictive models than the simple linear quantile regression estimates. This may also involve effective prediction of tertiary control activation. That would presumably improve the performance of the statistical arbitrage compared to our results. There are also other technological classes of players that could engage, eg storage operators or renewable generators with turn-down capabilities, but they would bring special operational constraints that would moderate, or possibly eliminate, their attractiveness. Overall, however, this research makes a strong contribution in demonstrating the operational benefit of opportunistic imbalancing and suggests that it should be positively viewed as a desirable step in market liberalization.

ACKNOWLEDGMENTS

The authors are grateful for constructive feedback from referees. Also Gregor Preisl and Oliver Albrecht advised on the model specification and participants at several conference presentations offered comments on earlier versions of this work including IAEE 2017 in Vienna and the EdF Portfolio and Risk Management Summit 2018 in Paris.

REFERENCES

APCS (2017). *Statistics*. February 2017. Retrieved from https://www.apcs.at/en/balancing-market/statistics/2015.

Austrian Power Grid (2016). *Rules regarding nomination of cross-border schedules*. February 2017.

Bertsch, J., C. Growitsch, S. Lorenczik, and S. Nagl (2016). "Flexibility in Europe's power sector—An additional requirement or an automatic complement?" *Energy Economics* 53: 118–131. https://doi.org/10.1016/j.eneco.2014.10.022.

Boogert, A. and D. Dupont. (2005). "On the effectiveness of the anti-gaming policy between the day-ahead and real-time electricity markets in The Netherlands." *Energy Economics* 27(5): 752–770. https://doi.org/10.1016/j.eneco.2005.03.006.

Brendan, C. (2016). "Dong Energy enters Demand-Side Response with Wind Power Balancing Service." *The Energyst*. Retrieved from https://theenergyst.com.

Browell, J. (2018). "Risk Constrained Trading Strategies for Stochastic Generation with a Single-Price Balancing Market." *Energies* 11(6). https://doi.org/10.3390/en11061345.

Davino, C. and M.V.D. Furno (2014). *Quantile Regression: Theory and Applications:* John Wiley and Sons. https://doi.org/10.1002/9781118752685.

Ding, H., P. Pinson, Z. Hu, J. Wang, and Y. Song (2017). "Optimal Offering and Operating Strategy for a Large Wind-Storage System as a Price Maker." *IEEE Transactions on Power Systems* 32(6): 4904–4913. https://doi.org/10.1109/TPWRS.2017.2681720.

Dowling, A.W., R. Kumar, and V.M. Zavala (2017). "A multi-scale optimization framework for electricity market participation." *Applied Energy* 190(Supplement C): 147–164. https://doi.org/10.1016/j.apenergy.2016.12.081.

Elexon. (2016). *Imbalance Pricing Guidance*. elexon.co.uk.

Elexon (2017). P305 *Post Implementation Review.* https://www.elexon.co.uk/wp-content/uploads/2017/02/P305-Post-Implementation-Review.pdf.

Elexon (2019a). *Investigations into Net Imbalance Volume Chasing in GB Electricity Market*. elexon.co.uk.

Elexon (2019b). *Market Index Definition Statement Review 2019*. elexon.co.uk.

Kaur, A., L. Nonnenmacher, H.T.C. Pedro, and C.F.M. Coimbra (2016). „Benefits of solar forecasting for energy imbalance markets." *Renewable Energy* 86: 819–830. https://doi.org/10.1016/j.renene.2015.09.011.

Kies, A., B. Schyska, and L. von Bremen (2016). "The Demand Side Management Potential to Balance a Highly Renewable European Power System." *Energies* 9(11): 955. https://doi.org/10.3390/en9110955.

Koenker, R. and G. Bassett (1978). "Regression Quantiles." *Econometrica* 46(1): 33–50. https://doi.org/10.2307/1913643.

Koenker, R.W., and V. D'Orey (1987). "Algorithm AS 229: Computing Regression Quantiles." *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 36(3): 383–393. https://doi.org/10.2307/2347802.

Koltsaklis, N.E., A.S. Dagoumas, and I.P. Panapakidis (2017). "Impact of the penetration of renewables on flexibility needs." *Energy Policy* 109: 360–369. https://doi.org/10.1016/j.enpol.2017.07.026.

Krishnamurthy, D., C. Uckun, Z. Zhou, P.R. Thimmapuram, and A. Botterud (2018). "Energy Storage Arbitrage Under Day-Ahead and Real-Time Price Uncertainty." *IEEE Transactions on Power Systems* 33(1): 84–93. https://doi.org/10.1109/TPWRS.2017.2685347.

Larsen, E.M., P. Pinson, J. Wang, Y. Ding, and J. Østergaard (2015). "The Cobweb Effect in Balancing Markets with Demand Response." *In Proceedings of European Electricity Market Conference 2015 IEEE, 2015,* 1–8.

Lisi, F. and E. Edoli (2018). "Analyzing and Forecasting Zonal Imbalance Signs in the Italian Electricity Market." *The Energy Journal* 39: 1–19. https://doi.org/10.5547/01956574.39.5.flis.

Möller, C., S.T. Rachev, and F.J. Fabozzi (2011). "Balancing energy strategies in electricity portfolio management." *Energy Economics* 33(1): 2–11. https://doi.org/10.1016/j.eneco.2010.04.004.

Nutaro, J. and V. Protopopescu (2009). "The Impact of Market Clearing Time and Price Signal Delay on the Stability of Electric Power Markets." *IEEE Transactions on Power Systems* 24(3): 1337–1345. https://doi.org/10.1109/TPWRS.2009.2021211.

OFGEM (2014). *Electricity Balancing Significant Code Review*.

Pinson, P., C. Chevallier, and G.N. Kariniotakis (2007). "Trading wind generation from short-term probabilistic forecasts of wind power." *IEEE Transactions on Power Systems* 22(3): 1148–1156. https://doi.org/10.1109/TPWRS.2007.901117.

van der Veen, R.A.C., A. Abbasy, and R.A. Hakvoort (2012). "Agent-based analysis of the impact of the imbalance pricing mechanism on market behavior in electricity balancing markets." *Energy Economics* 34(4): 874–881. https://doi.org/10.1016/j.eneco.2012.04.001.

van der Veen, R.A.C. and R.A. Hakvoort (2016). "The electricity balancing market: Exploring the design challenge." *Utilities Policy* 43: 186–194. https://doi.org/10.1016/j.jup.2016.10.008 .

Vandezande, L., L. Meeus, R. Belmans, M. Saguan, and J.M. Glachant (2010). "Well-functioning balancing markets: A prerequisite for wind power integration." *Energy Policy* 38(7): 3146–3154. https://doi.org/10.1016/j.enpol.2009.07.034.

Wang, Q., C. Zhang, Y. Ding, G. Xydis, J. Wang, and J. Østergaard (2015). "Review of real-time electricity markets for integrating Distributed Energy Resources and Demand Response." *Applied Energy* 138(Supplement C): 695–706. https://doi.org/10.1016/j.apenergy.2014.10.048.

Weber, C. and S. Just (2015). "Strategic Behavior in the German Balancing Energy Mechanism: Incentives, Evidence, Costs and Solutions." *Journal of Regulatory Economics* 48: 218–243. https://doi.org/10.1007/s11149-015-9270-6.

APPENDIX: QUANTILE ESTIMATION OF THE IMBALANCE DENSITY FUNCTIONS

Quantile regression (e.g. Davino and Furno, 2014) was used to model the imbalance densities as a function of the flow of market information and create quantile forecasts. The reason for using quantile regression, rather than for example GARCH-X, is that the shape of the densities, interpolated form all the estimated quantiles, is not restricted to a particular function and all the quantiles can depend separately upon the forecast wind and solar production at the time. Machine learning, eg via a neural network, could offer the same capability but without the specification transparency.

The formulation of a discrete set of conditional quantile functions can be solved based on the minimization problem proposed by (Koenker and Bassett 1978) for the q-th quantile regression, $0 < q < 1$, with a modification of the (Koenker and D'Orey 1987) algorithm. To estimate the winter and summer models (February and August), we used the 15min data for the respectively preceding months, January and July, 2015. Thus, to provide dynamic quantile estimates for $\widehat{imb}_{i,j}$, we investigated a plausible set of potential explanatory variables, from which we identified 3 generally significant regressors to describe the response variable for each discrete quantile estimate and imbalance settlement period. Furthermore, since we required different models to reflect the effects of different information time delays, the backshift j–r, with $r \in R\{1..8\}$, indicates the time-lagged structure for each decision-making context. The agent's estimation function for system imbalance was then specified as:

$$\widehat{imb}_{i,j} = \beta_{\mathrm{i1,j-r}} * imb_{j-r} + \beta_{\mathrm{i2,j-r}} \times wind_{j-r} + \beta_{i3,\mathrm{j-r}} * solar_{j-r} \times D_s + c_{\mathrm{i,j-r}}$$

being estimated separately for 10 discrete quantiles, indexed through *i,* in which:

imb_{j-r} is the system imbalance variable with a time lag of r,

$wind_{j-r}$ is the wind forecast error, calculated as the difference between the day ahead forecast and the latest value measured at (j-r), and

$solar_{j-r}$ is the solar forecast error, calculated as the difference between the day ahead forecast and the latest value measured at (j-r). Since the solar error was significant only in summer, there is a binary variable D_s which is 0 in winter and 1 in summer.

We used wind and solar day-ahead forecasts and outcome data from the Zentralanstalt für Meteorologie und Geodynamik (ZAMG). Forecast errors are calculated as differences between day-ahead forecasts and realized wind force and solar radiation. Forecast data is only available in hourly resolution whilst realization values are in 10 minutes resolution and so the values were linearly interpolated into 15 minutes intervals. Finally, from a discrete set of 10 quantile forecasts, we interpolated to create the discrete probability mass functions required for computing the optimal expected values as in Eq (12).