Well-functioning balancing markets: A prerequisite for wind power integration

Leen Vandezande[a,*], Leonardo Meeus[a], Ronnie Belmans[a], Marcelo Saguan[b] and Jean-Michel Glachant[b]

[a] K.U.Leuven, Electrical Engineering Department ESAT – ELECTA, Kasteelpark Arenberg 10 (PB 2445), B-3001 Heverlee, Belgium

[b] Université Paris-Sud XI, ADIS – GRM, 27 Avenue Ernest Lombart, 92260 Fontenay-aux-Roses, France

ABSTRACT

This article focuses on the design of balancing markets in Europe taking into account an increasing wind power penetration. In several European countries, wind generation is so far not burdened with full balancing responsibility. However, the more wind power penetration, the less bearable for the system not to allocate balancing costs to the responsible parties. Given the variability and limited predictability of wind generation, full balancing exposure is however only feasible conditionally to well-functioning balancing markets. On that account, recommendations ensuring an optimal balancing market design are formulated and their impact on wind generation is assessed. Taking market-based or cost-reflective imbalance prices as the main objective, it is advised that: (1) the imbalance settlement should not contain penalties or power exchange prices, (2) capacity payments should be allocated to imbalanced BRPs via an additive component in the imbalance price and (3) a cap should be imposed on the amount of reserves. Efficient implementation of the proposed market design may require balancing markets being integrated across borders.

The authors gratefully acknowledge the support and comments received from Michel Ceusters, James-Matthys Donnadieu, Bernard Malfliet, Patrick Mouttapa and Hans Vandenbroucke. The conclusions and remaining errors are the authors’ responsibility.

KEYWORDS

Electricity market design, balancing services, wind power integration

1. Introduction: Impact of increasing wind power penetration on balancing needs and costs

The presence of wind power in a system increases the need for regulating and reserve power in order to handle its variability and limited predictability. As indicated by various studies, balancing power requirements are expected to increase proportionally with growing wind power penetration (VTT Technical Research Center of Finland, 2007).

The precise impact of an increasing amount of wind power capacity on balancing needs however depends on multiple factors: geographical aspects including the size of the control area and the geographical spread and aggregation of wind farms, initial load variations and their similarity to wind power fluctuations, operational routines of the power system and the prevailing market architecture – e.g. the frequency of load and wind power forecast updating and thereupon based re-scheduling (depending on day-ahead/intra-day gate closure times) –, the quality and accuracy of the wind power forecast system used, etc. (EWEA and Tradewind, 2009).

Furthermore, additional requirements differ depending on the type of balancing services. In the primary control time scale, the effect of wind power variations on the system operation is very small even at considerable penetration (Holltinen, 2004). Moreover, the amount of primary reserves in both UCTE and NORDEL is determined on a synchronous area-wide basis, implying that an increase in wind power capacity in a control area will not directly affect[1] its need for primary reserves. Only the saturation level of its primary reserves, i.e. the band of primary reserves being continuously activated, may be raised. As primary control services are typically either unpaid or remunerated for capacity only (€/MW) – and not for energy (€/MWh) – the latter will however have no impact on total balancing costs. Therefore, the optimal market design for this type of service is not extensively considered in this article. The main focus is by contrast on secondary and tertiary control services (UCTE) and fast reserves (NORDEL), i.e. the services activated immediately after the primary control time scale[2], as wind power variations are mainly dealt with by them and their costs – i.e. a remuneration for energy (€/MWh) and sometimes capacity (€/MW) – consequently rise with growing wind power penetration.

Faced with these increased balancing needs and costs the question arises who should be held responsible for them. In several European countries, wind generation is so far not fully exposed to market risks including balancing responsibility. While such an approach can be justified in countries with a low share of wind in the total generation mix, countries with a higher wind share and a significant impact of wind on their system should burden more responsibility and risks on wind generation in order to give them an incentive for cost-reflective market behaviour and as such limit the indirect costs to society. Equal balancing rules for all market participants may for instance encourage wind power producers to provide more accurate forecasts of their generation – in order to reduce their own balancing costs – which may in turn lead to an increased system balancing efficiency. For an extensive discussion on this and other positive effects of full market participation of wind energy, see Hiroux and Saguan (2009). As illustrated in Klessman et al. (2008), full balancing risk exposure is however only feasible conditionally to well-functioning intra-day and balancing markets.

Wind power producers can resume their balance responsibility in several ways. Like other Balance Responsible Parties (BRP), they can participate in wholesale markets to fine-tune their generation portfolio up to the last gate closure – indicating the importance of liquid intra-day markets with gate closure near the real-time – and rely on the Transmission System Operator (TSO) to resolve their remaining imbalances. For a discussion on intra-day market design and measures to increase its liquidity, see Weber (2009). Depending on the imbalances incurred, an imbalance charge (€/MWh) is imposed by the TSO – pointing out the importance of efficient balancing markets with reliable imbalance prices. A second option consists in holding own regulating power units or electricity storage and in that way avoid imbalance costs. However, this implies additional investments and operation costs to be considered when projecting a new wind farm. Similarly, contracts could be concluded with interruptible loads. Finally, wind power producers can chose to outsource their balance responsibility to another market party (with conventional generation units) at a certain cost. This option limits their exposure to balancing risks, implying a less risky business case when projecting a new wind farm.

This article focuses on the design of balancing markets in Europe taking into account an increasing wind power penetration. Section 2 and 3 discuss design options with respect to the allocation of energy and capacity costs respectively and formulate recommendations ensuring market-based or cost-reflective imbalance prices. The impact of these recommendations on wind generation is also assessed. Section 4 considers the implications of implementing the – in the previous sections – proposed market-based design.

2. Ensuring cost-reflective imbalance prices: allocation of energy payments

Balancing markets provide market parties with a 'last resort' for energy transactions. The real-time or imbalance prices expected to be brought forth by this market are reflected in wholesale prices and consequently affect market parties' decisions at the forward stage. For this reason, electricity markets can only function efficiently conditionally to market-based or cost-reflective imbalance prices (Tractebel Engineering and K.U.Leuven, 2009). The need for imbalance prices reflecting the true costs imbalanced BRPs incur to the system becomes even more important with increasing wind power penetration as the variability and limited predictability of wind generation prevents them – more than conventional generation – from being perfectly balanced.

Imbalance prices are market-based insofar as they fully reflect *all* procurement expenses incurred by the TSO for delivering energy in real-time. As such, imbalance prices should, in principle, correctly pass on both energy (€/MWh) and – if applicable – capacity payments (€/MW). While the former are payments on a settlement period basis for the actual real-time delivery of regulating power, the latter are payments made beforehand for holding reserve capacity available. Design options with respect to the allocation of energy payments are considered below; the optimal allocation of capacity payments is discussed in Section 3.

2.1. One versus two price systems

Imbalance prices are usually based on up- and downward regulating power offers accepted by the TSO for real-time balancing. They are based on either the price of the *marginally* accepted up- or downward regulating offer or the *average* price of all accepted up- or downward regulating offers, depending on how Balancing Service Providers (BSP) are remunerated. For a discussion on the pros and cons of remunerating BSPs by means of marginal pricing versus average or pay-as-bid pricing, see for instance Littlechild (2007). Briefly, there is a widely held view that marginal pricing is economically more correct and will lead to a more efficient allocation of resources than average pricing. Apart from the choice between marginal and average pricing, a difference also exists between single and double imbalance pricing schemes. Table 1 and Table 2 represent a typical one and two price system.

Under a single imbalance pricing scheme or one price system, imbalance prices correspond to the marginal procurement price of balancing services, i.e. either upward or downward regulating services depending on the overall status of the system. The same imbalance price – though with a different sign – is applied for remaining short and long positions[3], making the imbalance settlement theoretically a zero sum game for the TSO.

Table 1: Imbalance settlement through a typical one-price system[a]

Under a double imbalance pricing scheme or two price system on the contrary, a different imbalance price is applied for positive and negative BRP imbalances. While BRP imbalances contributing to the system imbalance are settled at prices based on the – usually average – procurement costs of balancing services, BRP imbalances counteracting the system imbalance are settled on the basis of wholesale price indices, typically power exchange prices. Compared to a one price system, under which settlement of BRP imbalances opposing the system imbalance is based on marginal costs – i.e. the additional cost the TSO would have incurred if the BRP concerned was not imbalanced –, the latter is often implemented to avoid generators speculating on the direction of the system imbalance – i.e. creating a short position if they expect the system imbalance to be long and vice versa. However, it is rather doubtful whether generators would change their position on the basis of a – very short-term – settlement period. Such “gaming” actions are unlikely to be profitable for generators and endangering the system security at the same time or, in other words, profitable actions will normally go hand in hand with actions increasing the system security.

Given the presence of power exchange prices (and possibly penalties – cf. below), a two-price system no longer implies a zero-sum game for the TSO, which should not have financial interest in the imbalance settlement. Accordingly, insofar the difference is not used by the TSO to cover other costs in real time (e.g. staffing and IT costs), it should result in a reduction of transmission tariffs. But even if this is done, it still entails a transfer of money from inflexible users – such as wind generators – to average users. Furthermore, a two-price system puts small market parties – again often including wind generators – at a disadvantage as it involves lower imbalance costs for larger market parties due to netting. For that reason, small market parties are ‘gently forced’ to outsource their balance responsibility. On the contrary, under a one-price system, no extra discrimination is made according to the size of market participants.

Table 2: Imbalance settlement through a typical two-price system[b]

Finally, a two-price system sometimes includes a multiplicative component or so-called penalty that affects BRPs with regard to their position before real-time. This penalty typically affects negative imbalances more than positive ones, thus encouraging BRPs to avoid short positions. Other than for BRP motivation to be balanced – and associated security safeguarding – penalties are imposed for practical reasons such as accounting – for instance to generate extra revenues for the recovery of intra-settlement period imbalances – and the recovery of capacity payments (cf. infra).

Insofar as they are not cost-reflective, penalties can give rise to undesirable BRP behaviour, including over-contracting in the wholesale market, withholding services for own use and nominating less than the expected injections. These negative side-effects are more extensively discussed below using some basic examples. Note that these examples have been kept simple for purposes of clarification and do not aim to be exhaustive. More specifically, the following assumptions are made:

- Both the one- and two-price system are based on marginal procurement prices (MP).
- Marginal procurement prices are expressed as a percentage of the day-ahead price: e.g. $MP_u = 1.5*P_{DA}$ and $MP_d = 0.5*P_{DA}$. While marginal prices for upward regulation are higher than day-ahead prices, marginal prices for downward regulation are lower. Note that, in practice, the supply curve for balancing services is typically not linear. In thermal power systems – contrary to power systems with substantial amounts of hydro and/or wind power (e.g. Nordel) – downward regulation is usually relatively easier and these services are consequently cheaper – i.e. their marginal price deviates less from the day-ahead price compared to upward-regulating services. Because of this, BRPs in such systems already exhibit a natural tendency to strive for long rather than balanced positions[4].
- For the moment, marginal procurement prices for upward regulation are usually higher than day-ahead prices. However, the better markets continue to function – and the more arbitrage opportunities are exploited – the more day-ahead and imbalance prices will converge. Note, however, that even if both prices are equal, BRPs would still make a difference between buying

energy on the wholesale market or the balancing market. They would rather buy wholesale to hedge against typically higher and more volatile imbalance prices. This is the case as not all generation resources can be controlled fast enough to deliver energy in the real-time.

- For simplification, the day-ahead price equals 1 ($P_{DA} = 1$).
- Penalties under the two-price system are higher for short positions than for long ones ('asymmetric' penalties): $\text{Penalty}_u = 0.4$ and $\text{Penalty}_d = 0.25$.
- BRPs are unaware of the system imbalance: 50% of time positive/negative

In Table 3 and Table 4 the above assumptions are applied to the one- and two-price system respectively.

Table 3: Input data for examples – One-price system

Table 4: Input data for examples – Two-price system

Impact of imbalance pricing on wholesale trade

To illustrate the potential impact of a two-price system with asymmetric penalties on wholesale markets, assume a BRP – owning only load – with an expected load of 100 MW or, more specifically, a load of 90 MW or 110 MW, each for 50% of the time. As calculated in Table 5 under a one-price system, the BRP has no preference between buying energy on the wholesale market or the balancing market.

For instance, if the BRP procures 90 MW on the day-ahead market – i.e. less than the expected load – it only pays 90 beforehand. In real-time, the BRP is balanced during half of the time. The rest of the time, it faces a negative imbalance of -20 MW for which it pays an imbalance charge to the TSO. This charge is calculated on the basis of an imbalance price equalling 1.5 or 0.5, depending on the direction of the system imbalance. Expected total costs for the BRP come to 100. If the BRP procures 100 MW on the day-ahead market, equalling the expected load, it pays 100 beforehand. In real-time, it is faced with negative and positive imbalances of -10 and +10, each for 50% of the time. It accordingly pays and receives similar imbalance charges and its expected total costs are again 100. If the BRP procures 110 MW on the day-ahead market – exceeding the expected load – it pays 110 beforehand. In real-time, it is balanced again for half of the time, for the rest being subject to a positive imbalance of 20 for which it receives an imbalance charge from the TSO. This charge is calculated on the basis of an imbalance price equalling 1.5 or 0.5, depending on the direction of the system imbalance. The expected final outcome is the same as in the other two cases.

Table 5: Example on the impact of imbalance pricing on wholesale trade

However, under a two-price system with asymmetric penalties, the BRP is inclined to over-contract energy on the wholesale market and thus avoid a short position. For instance, if the BRP procures 90 MW on the day-ahead market – i.e. below the expected load – it pays only 90 beforehand. In real-time, the BRP is balanced for half of the time. The rest of the time, it faces a negative imbalance of -20. It therefore pays an imbalance charge to the TSO, calculated on the basis of an imbalance price equalling 2.1 or 1, depending on the direction of the system imbalance. Its expected total costs are 105.5. If the BRP procures 100 MW on the day-ahead market – equalling the expected load – it pays 100 beforehand. In real-time, it is faced with negative and positive imbalances of -10 and +10, each for 50% of the time. Since the penalty imposed on short positions is higher, the imbalance charge the BRP pays to the TSO is higher than that which it receives. Its expected total costs are 104.25. If the BRP procures 110 MW on the day-ahead market – exceeding the expected load – it pays 110 beforehand. In real-time, it is balanced for half of the time. The rest of the time, it has a positive imbalance of 20 and therefore receives an imbalance charge from the TSO. This is calculated on the basis of an imbalance price equalling 1 or 0.4, depending on the direction of the system imbalance. The BRP's expected total costs are 103.

A comparison of the expected costs under a two-price system with asymmetric penalties shows that the BRP would naturally prefer to increase its day-ahead purchases as a hedge against real-time short positions and the associated higher penalties. This BRP behaviour has the overall effect of increasing wholesale prices, as modelled in Saguan (2007) and Saguan and Glachant (2007). Given that BRPs – at least in thermal power systems – already exhibit a natural tendency to strive for long rather than balanced positions – because regulating downward is easier than regulating upward and/or downward regulating services are cheaper than upward regulating services –, this behaviour should not be reinforced through the introduction of penalties.

Impact of imbalance settlement on the provision of balancing services (upward regulation)

To illustrate the potential impact of a two-price system with asymmetric penalties on balancing services supply, assume a BRP – owning both generation and load – with generation of 110 MW and an expected load of 100 MW or, more specifically, a load of 90 MW or 110 MW, each for 50% of the time. As calculated in Table 6, under a one-price system, the BRP has no preference between offering balancing services to the TSO via the balancing market or keeping services for own use. Note that in this example, the activation cost of balancing services is taken into account. This activation cost is assumed to be equal to the marginal procurement price of upward regulating services, being 1.5.

For instance, if the BRP offers 10 MW to the TSO, its services have a maximal 50% chance of being activated in real-time – given a negative system imbalance for half of the time. It therefore receives a remuneration based on the marginal price for upward regulation, i.e. 1.5, which exactly compensates it for the activation cost. Furthermore, the BRP is exposed to negative and positive imbalances of -10 and +10, each for 50% of the time. The BRP pays and receives similar imbalance charges accordingly. Its expected final income is 0. However, if the BRP keeps its 10 MW for its own use, it can avoid short positions in real-time. It will however only activate its services on the condition that the imbalance charge for short positions exceeds the activation cost[5]. With imbalance prices of 1.5 and 0.5 – depending on the system imbalance – and an activation cost of 1.5, the BRP will never activate its 10 MW and will prefer to be short instead. As a result, the BRP is exposed to negative and positive imbalances of -10 and +10, each for 50% of the time. It accordingly pays and receives similar imbalance charges and its expected final income is again 0.

Table 6: Example on the impact of imbalance pricing on the provision of balancing services

However, under a two-price system with asymmetric penalties, the BRP is inclined to keep its excess generation for its own use, thus avoiding a short position if the load is higher than expected and the imbalance charge for short positions exceeds the activation cost. For instance, if the BRP offers 10 MW to the TSO, its services have a maximal 50% chance of being activated in real-time – given a negative system imbalance for half of the time. It therefore receives a remuneration based on the marginal price for upward regulation, i.e. 1.5, which exactly compensates it for the activation cost. Furthermore, the BRP is exposed to negative and positive imbalances of -10 and +10, each for 50% of the time. Since the penalty imposed on short positions is higher, the imbalance charge the BRP pays to the TSO is higher than that which it receives. Its expected final income amounts to -4.25. However, if the BRP keeps its 10 MW for its own use, it can avoid short positions in real-time. It will only activate its services on the condition that the imbalance charge for short positions exceeds the activation cost. With imbalance prices of 2.1 and 1 – depending on the system imbalance – and an activation cost of 1.5, the BRP will only activate its 10 MW if the former imbalance price holds. As a result, the BRP is exposed to negative and positive imbalances of -10 and +10 for 25% and 50% of the time respectively. It pays and receives imbalance charges accordingly. By activating its services for 25% of the time, the BRP can partly avoid paying the relatively higher imbalance charge for short positions. Its expected final income is 2.75.

A comparison of both outcomes indicates that the BRP will prefer to keep any excess generation for its own use as a hedge against real-time short positions and the associated penalties. This 'self-regulating' behaviour has a negative effect on the supply of energy in the real-time market and consequently limits the ability of TSOs to balance the system. This effect has been mentioned in Newbery and McDaniel (2002) and Cornwall (2001). In extreme cases, this behaviour could result in each BRP holding a back-up for its own largest plant, which is, of course, highly inefficient.

Impact of imbalance settlement on nominations

Note that this example is similar to the first one.

To illustrate the potential impact of a two-price system with asymmetric penalties on the accuracy of nominations – or, differently said, on the reliability of BRP bids submitted to the market and the associated scheduled portfolio declared to the TSO –, assume a BRP – consisting of only generation - with an expected generation of 100 MW or, more specifically, a generation equalling 90 MW or 110 MW, each during 50% of the time. As calculated in Table 7, under a one-price system, the BRP has no preference between nominating according to or different from its expected level of generation.

For instance, if the BRP sells 90 MW– i.e. less than its expected generation level – on the day-ahead market, it receives only 90 beforehand. In real-time, it is balanced for half of the time. The rest of the time, it faces a positive imbalance of +20. The BRP therefore receives an imbalance charge from the TSO. This charge is calculated on the basis of an imbalance price equalling 1.5 or 0.5, depending on the direction of the system imbalance. The BRP's expected total income is 100. If the BRP sells 100

MW– equalling the expected generation level – on the day-ahead market, it receives 100 beforehand. In real-time, it is faced with negative and positive imbalances of -10 and +10, each for 50% of the time. It pays and receives similar imbalance charges accordingly and its expected total profit is again 100. If the BRP sells 110 MW– exceeding the expected generation level – on the day-ahead market, it receives 110 beforehand. In real-time, it is balanced for half of the time. For the rest of the time, it has a negative imbalance of -20. It therefore pays an imbalance charge to the TSO. This is calculated on the basis of an imbalance price equalling 1.5 or 0.5, depending on the direction of the system imbalance. The BRP's expected final outcome is the same as in the other two cases.

Table 7: Example on the impact of imbalance pricing on nominations

However, under a two-price system with asymmetric penalties, the BRP is inclined to nominate less than its expected generation level, thus avoiding a short position. For instance, if the BRP sells 90 MW– i.e. less than the expected generation level – on the day-ahead market, it receives only 90 beforehand. In real-time, it is balanced for half of the time. For the rest of the time, it faces a positive imbalance of +20. The BRP therefore receives an imbalance charge from the TSO. This charge is calculated on the basis of an imbalance price equalling 2.1 or 1, depending on the direction of the system imbalance. The BRP's expected total income is 97. If the BRP procures 100 MW – equalling the expected generation – on the day-ahead market, he receives 100 beforehand. In real-time, it is faced with negative and positive imbalances of -10 and +10, each for 50% of the time. Since the penalty imposed on short positions is higher, the imbalance charge the BRP pays to the TSO is higher than that which it receives. The BRP's expected total profit is 95.75. If the BRP sells 110 MW– exceeding the expected generation level – on the day-ahead market, it receives 110 beforehand. In real-time, it is balanced for half of the time. The rest of the time, it faces a negative imbalance of 20. The BRP therefore pays an imbalance charge to the TSO. This is calculated on the basis of an imbalance price equalling 1 or 0.4, depending on the direction of the system imbalance. The BRP's expected final outcome is 94.

A comparison of the profits expected under a two-price system with asymmetric penalties indicates that the BRP will prefer to under-nominate its expected injections as a hedge against real-time short positions and the associated penalties. This behaviour has a negative effect on the reliability of the information TSOs receive through the nomination process.

2.2. Separate imbalance settlement to counteract the side effects of two-price systems

The above section discussed potential negative side effects of a two-price system assuming a one-step imbalance volume calculation or, in other words, a single settlement for generation and load. However, some European countries settle generation and load separately. Depending on its implementation, a separate imbalance settlement can partly counteract the negative side effects of a two-price system with asymmetric penalties.

For instance, the implementation of a harmonised imbalance settlement in the Nordic region in early 2009 – which settles generation using a two-price system and load using a one-price system – may have the following effects (Nordel, 2007 and NordREG, 2008):

- TSO gains under the two-price system – which should be redistributed by reducing transmission tariffs – result in a transfer from average to inflexible users rather than the other way around.
- Small market participants owning only load will not be discriminated against compared to larger ones as generation and load are settled separately, the latter according to a one-price system.
- BRPs owning only load will be less inclined to over-contract in the wholesale market as settlement occurs on the basis of a one-price system.
- BRPs owning both generation and load will have less incentives to hold back services for own use as generation and load are settled separately.
- BRPs owning only generation will still have a tendency to nominate less than their expected injections as settlement occurs on the basis of a two-price system. This puts small generators owning only few units – like wind power producers – at a disadvantage.

2.3. Recommendation ensuring an optimal market design and its impact on wind generation

Following the analyses in the previous sections, it is recommended to avoid the use of non-market based components such as power exchange prices and penalties in the imbalance settlement. Note that the implementation of a separate imbalance settlement for generation (two-price system) and load (one-price system) like recently implemented in the Nordic countries has its merits but does not completely counteract the negative side effects of asymmetric penalties.

Implementation of this recommendation positively affects wind generation in three ways. First, an abolition of penalties has a reductive effect on overall imbalance prices, as for instance illustrated in Helander et al. (2008) for wind generation. Secondly, given that the correlation between system imbalance and individual wind power imbalances rises with increasing wind penetration level, wind generation's imbalance costs are lower under an imbalance settlement without penalties – payable only by those BRPs aggravating the system imbalance (Weber, 2009). Third, strategies – as illustrated in the previous sections – to avoid short real-time positions in case of penalties are often more easily executable and profitable for conventional generation resources than for wind generation.

The above analyses and recommendation fit with the advise formulated in Barth (2008) and Wilmar (2005) to implement a one-price system – based on marginal prices – rather than a two-price system. The former system is preferred as it provides BRPs with adequate scarcity signals, contrary to the latter in which imbalances in the opposite direction of the system imbalance are not rewarded. Note that this design choice implies that wind power producers – and more in general all BRPs – are exposed to extreme imbalance prices in case flexible generation is scarce. However, such high prices usually do not occur without any announcement as signals are provided by the wholesale markets organised beforehand.

3. Ensuring cost-reflective imbalance prices: allocation of capacity payments

Contrary to primary control services, it would be preferable to avoid remunerating secondary and tertiary control services (UCTE) and fast reserves (NORDEL) for capacity, amongst others because of the difficulties in accurately allocating the associated costs (cf. infra). However, three fundamental arguments account for the use of capacity payments and explain why this type of remuneration is currently implemented in some countries.

Firstly, balancing markets typically exhibit higher but more volatile prices and smaller volumes as activation is dependent on the system state. Consequently, revenues are often more volatile than in wholesale markets, inciting generators to sell on the wholesale rather than the balancing market. In such case, capacity payments – yielding a guaranteed income – can serve as a risk premium to attract more BSPs. Secondly, balancing markets – and in general all electricity markets – exhibit non-convexities, such as start-up costs and minimum output levels. To ensure efficient dispatch in the presence of non-convexities and simultaneously safeguard uniform or marginal real-time energy procurement prices, an additional capacity payment can be useful, especially in small markets where e.g. start-up costs have a relatively larger impact. Thirdly, balancing markets in several countries still exhibit regulated prices for real-time energy[6], making it impossible for BRPs to pass on all of their costs via their real-time energy bids (including opportunity costs and actual costs for keeping services online). Capacity payments are thus a means of recovering these remaining costs. However, if there is a well-functioning and unrestricted balancing market, BSPs have the opportunity to pass on all costs via the real-time energy price. In view of these arguments, it is understandable that some countries currently remunerate services other than primary control and disturbance reserves for capacity. However, this type of remuneration should only be transitional and should preferably be phased out.

3.1. Socialisation among grid users or BRPs versus additive component in the imbalance price

Contrary to energy payments, capacity is procured for a time period far exceeding the settlement period. Consequently, its associated costs cannot be directly attributed to imbalanced BRPs. A choice should therefore be made between one of the following cost allocation methods.

Socialisation among grid users via transmission tariffs

Socialisation of capacity payments among grid users (consumers and producers) does not entail cost-reflective imbalance prices: the resultant imbalance prices are too low as they do not include all procurement costs. Consequently, BRPs are given fewer incentives to balance their portfolio using wholesale markets (day-ahead and intra-day markets) and increasingly rely on the balancing market.

However, for services which are mainly deployed for capacity purposes and remunerated for capacity only (i.e. primary control services and sometimes disturbance reserves), socialisation of capacity payments is justified. As these services mostly operate as a kind of public security insurance and/or are needed even if all BRPs are balanced, their costs should not be allocated to individual BRPs according to the extent of their imbalance. With a view to avoiding over-contracting by TSOs and protecting grid users against excessive transmission tariffs, the amount of capacity payments for this type of services should be regulated.

There are currently many methods of socialising capacity costs among grid users. Typically countries pass through the costs of primary reserves on grid users. Many countries do however allocate the costs of other services – deployed for real-time energy delivery rather than capacity purposes – via the transmission tariffs as well.

Socialisation among BRPs via periodical fee

Although socialisation of capacity payments among BRPs via a periodical fee is already an improvement on socialisation via transmission tariffs, it does not yet provide BRPs with the right incentives. Since the periodical fee is fixed (€/period) or proportional to BRPs' injections (generation – import – purchases) or off-takes (consumption – export –sales) (€/MWh of injections/off-takes) – i.e. the BRPs' size – rather than proportional to BRP's imbalances, imbalance prices will again be too low, encouraging BRPs to be over-reliant on the balancing market.

It is important to note that countries implementing a pure one-price system can only pass on energy costs via the imbalance price and have no choice but to allocate capacity costs via a socialisation among grid users or BRPs. Consequently, pure one-price systems – like two-price systems with non-market based components – are not fully cost-reflective.

Here are some examples of current practices of a socialisation among BRPs:

- In France, a monthly fee – the 'prix proportionnel au soutirage physique' – is imposed on BRPs proportional to their off-takes to recover capacity payments of the 'réserves rapides' (reserves with an activation time of 15 min.).
- In the United Kingdom, capacity payments are partially allocated to BRPs via 'BSUoS charges', a fee imposed per settlement period (1/2 hour) proportional to BRPs' injections or off-takes.
- In the Nordic countries, the harmonised imbalance settlement implemented since early 2009 partially allocates capacity costs through a monthly fixed fee and a fee proportional to BRPs' measured generation or consumption.

Allocation to BRPs proportional to their imbalances via additive component in imbalance price

The third and most cost-reflective method consists in the inclusion of capacity costs in the imbalance price (€/MWh of imbalances). Such allocation of capacity payments is similar to the allocation of fixed costs under *Ramsey-Boiteux* pricing, whereby fixed costs are recouped from customers by charging them prices in excess of marginal costs, in inverse proportion to their demand elasticity. The argument of Ramsey-Boiteux pricing has been used similarly by Hogan (2006) with respect to the allocation of so-called “Resource Sufficiency Costs” in Midwest USA (MISO).

Based on this, capacity payments can be recovered by means of an additive component[7] (component$_{cap}$) on top of the marginal procurement price of upward or downward regulating services (MP$_{u/d}$). In this case, inelastic customers include all the BRPs that ‘chose’ to be imbalanced despite the imbalance price being higher than the marginal cost of upward or downward regulation. Wind generators usually belong to the this category. Allocation of both energy and capacity payments through the real-time energy price is summarised in Table 8.

Table 8: Allocation of capacity payments via the imbalance price[c]

Note that the resulting imbalance pricing system exhibits characteristics of both a one-price and a two-price system. It is similar to a one-price system in that it allocates energy costs using marginal procurement prices only. It is also similar to a two-price system in that it entails different imbalance prices depending on the sign of the BRP’s imbalance, but – contrary to a two price system – it does not include non-market based components.

To ensure a cost-reflective imbalance price, it is vital to determine the additive component accurately. Spreading out of capacity payments over all imbalanced BRPs during the time period of capacity reservation, the additive component can only be calculated using historical figures on the amount and extent of BRP imbalances. Consequently, an exact recovery of capacity payments using the additive component is unattainable. Moreover, the longer the terms of capacity reservation, the less accurate the additive component will be. Therefore, from a cost allocation point of view, capacities are preferably procured on a short-term basis, e.g. daily rather than yearly capacity payments. Short capacity reservation periods also involve a fast learning curve with respect to the necessary amount of reserves, making capacity payments a more ‘controllable’ cost. However, the impact of shorter reservation periods on competition is uncertain. On the one hand, short-term capacity payments reduce market foreclosure, but on the other hand, they might provide incumbents with the opportunity to game on a more regular basis. For this reason, the optimal length of the reservation period should be defined taking into account the impact on both cost allocation and competition. The preferences of

BSPs and TSOs should also be considered. It is likely that they would both prefer longer reservation periods as this reduces the risks they face.

Here are some examples of current practices as regards additive components in the imbalance price:

- In the Nordic countries, the harmonised imbalance settlement proposal to be implemented in early 2009 provides for a volume fee on consumption imbalances to recover part of the capacity payments.
- In Austria, the imbalance settlement system applied since 2006 allocates capacity costs through a component included in the real-time energy price that gradually increases proportional to the extent of the system imbalance during the settlement period concerned.

Note that in the latter case, the additive component has been implemented in such way that it provides additional – but redundant or even wrong – incentives for BRPs, which should of course be avoided; the gradual increase in the additional component is achieved through the addition of a non-market based component acting as a kind of ‘security penalty’.

3.2. Recommendation ensuring an optimal market design and its impact on wind generation

Following the analysis in the previous section, it is recommended to use an additive component in the imbalance price to allocate capacity payments for services delivering a significant amount of real-time energy. Socialisation of these capacity payments among grid users or BRPs should be discouraged as it results in relatively too low imbalance prices and a potential over-reliance of BRPs on the balancing market.

Implementation of this recommendation may however have a negative impact on wind generation. The additive component mainly affects inelastic consumers like wind generators – given their limited predictability and variability – that have no other possibility than relying on the balancing market for their last resort energy needs. As such, wind generators are very likely to bear a significant part of the capacity payments carried out by the TSO, having an uplifting effect on their imbalance costs. However, these capacity costs are also partly caused by wind and – as indicated in Section 1 – the more wind power penetration, the less bearable for the system not to allocate these “hidden” costs to the responsible parties. Furthermore, implementation of a third recommendation capping the amount of reserves (cf. infra) dampens the increasing impact on imbalance prices.

In a nutshell, while from a wind generator perspective socialisation of capacity payments among all grid users is most likely preferred above an allocation among imbalanced BRPs only, the latter option is best-suited to ensure a market-based or cost-reflective design.

4. Implications of implementing a market-based balancing design

The implementation of a market-based balancing design as put forward in Section 2 and 3 has two major implications.

4.1. Need for restrictions on the amount of reserves

Allocation of capacity payments via an additive component has a negative impact on new entrants – like wind generators – rather than incumbents, the former being the most inelastic customers since they usually do not have the opportunity to balance their own portfolios. Reservation of balancing services mainly deployed for real-time energy delivery should therefore be kept to a minimum[8]. Consequently, in order to avoid barriers to entry, a cap should be imposed on the amount of reserves so that the share of component_{cap} in the final imbalance price is small compared to the marginal upward or downward regulation price. As a rule of thumb, reservations of services should only be accepted when needed to compensate for the higher revenue volatility in balancing markets compared to wholesale markets. The appropriateness of the level of the imposed cap can be verified by monitoring whether (1) the real-time energy delivery of the reserves concerned is marginal in comparison to the total delivered real-time energy and (2) the additive component has only a marginal effect on the imbalance price.

The relevance and necessity of a regulated amount of reserves is reinforced by the fact that some TSOs are currently considering substantially increasing the amount of reserves or even building their own plants (or leasing or taking over old plants) – this being an extreme form of capacity payments – to ensure sufficient availability of services for real-time energy delivery purposes. A lack of confidence in the balancing market and the associated fear of a shortfall in reserves are often the reasons behind such plans.

However, over-contracting reserves gives rise to several negative side-effects[9]:

- It limits trade opportunities on the wholesale market, thus increasing price differences between wholesale and balancing markets.

- It reduces real-time energy prices – even when these should be high because of generation capacity scarcity – which could eventually result in the disappearance of the balancing market.
- It might increase moral hazards by giving BRPs an implicit guarantee that all imbalances can be covered by reserves procured by the TSO.

4.2. Infeasibility of market-based design on national level

Currently implemented balancing market designs across Europe often deviate significantly from the market-based design proposed above. However, these deviations are understandable in a national context, considering market concentration and the (non)-existence of a well-functioning intra-day market.

As indicated in the Energy Sector Inquiry (EC, 2007), balancing markets in most Member States are currently highly concentrated, mirroring the concentration levels in generation in many wholesale markets. Concentration in balancing is even higher due to the fact that not all generators can supply regulating power in view of the technical criteria. This concentration simply does not allow some balancing markets to function properly on a national scale. This explains why many balancing 'markets' are currently more regulated than market-based according to the interpretation given above.

The potential infeasibility of a market-based design on a national scale implies that cross-border balancing should be implemented first and balancing market designs should only be harmonised afterwards. As illustrated in Tractebel Engineering and K.U.Leuven (2009), cross-border balancing can be implemented without having to meet unrealistic or overly expensive preconditions.

Note that the implementation of cross-border balancing would be beneficial for wind generation as well; it would not only provide access to cheaper regulating resources – implying lower imbalance prices – but the creation of larger control areas may also enable an extensive aggregation of wind farms to smooth imbalances.

5. Conclusions

Most policy analyses on wind power integration mainly compare RES-E support policies such as feed-in tariffs and quota systems. However, the success of wind energy deployment does not only depend on the efficiency of support policies but also on the design of electricity markets and the interaction between both (Ragwitz et al., 2007 and Hiroux and Saguan, 2009). This article therefore discussed the optimal design of balancing markets in Europe taking into account an increasing wind power penetration.

Due to its variability and limited predictability, the presence of wind power in a system increases balancing power requirements. Faced with these increased balancing needs and costs the question arises who should be held responsible for them. In several European countries, wind generation is so far not burdened with full balancing responsibility. However, the more wind power penetration, the less bearable for the system not to allocate balancing costs to the responsible parties. Full balancing risk exposure is however only feasible conditionally to well-functioning intra-day and balancing markets.

Optimally designed balancing markets should bring forth reliable imbalance prices reflecting all costs imbalanced BRPs incur to the system. Taking market-based or cost-reflective imbalance prices as the main objective, it has been recommended in this article that: (1) the imbalance settlement should not contain penalties or power exchange prices, (2) capacity payments should be allocated to imbalanced BRPs via an additive component in the imbalance price and (3) a cap should be imposed on the amount of reserves. While the first recommendation was shown to have a multiple reductive effect on imbalance costs of wind generation, the second one might have a negative impact. As the additive component mainly affects inelastic consumers like wind generators, they are very likely to bear a significant part of the capacity payments carried out by the TSO, having an uplifting effect on their imbalance costs. However, these “hidden” capacity costs are also partly caused by wind. Furthermore, implementation of the third recommendation capping the amount of reserves dampens the increasing impact on imbalance prices.

Efficient implementation of the in this article proposed market design may require balancing markets being integrated across borders. In other words, a changeover to full balancing responsibility across Europe should not only be preceded by harmonisation of balancing market designs in accordance with the market-based or cost-reflective design but also by implementation of cross-border balancing. In addition, this changeover cannot be considered separately from a harmonisation of support mechanisms.

References

Barth R., Weber C., Swider D.J., 2008. Distribution of costs induced by the integration of RES-E power, Energy Policy, 36(8), August 2008, Pages 3107-3115

Cornwall, 2001. NETA – Is the glass half empty or half full?, Consulting report, Available at http://www.hks.harvard.edu/

European Commission (EC), 2007. DG Competition report on Energy Sector Inquiry, 10 January 2007, Available at http://ec.europa.eu/competition/

European Wind Energy Association (EWEA), 2009. The Economics of Wind Energy, March 2009, Available at http://www.ewea.org

Finon D., Meunier G. and Pignon V. (2008) The social efficiency of long-term capacity reserve mechanisms, Utilities Policy 16, pages 202-214

Helander A.,Holttinen H., Paatero J., 2008. Impact of wind power on the power system imbalances in Finland, Impact of wind power on the power system imbalances in Finland, 7th International Workshop on Large-Scale Integration of Wind Power into Power Systems and on Transmission Networks for Offshore Wind Farms, Madrid, Spain, 26 - 27 May 2008., Pages 183-187, Energynautics

Hiroux C., Saguan M., 2009. Large-scale wind power in European electricity markets: time for revisiting support schemes and market designs?, this Energy Policy issue

Hogan W. W. (2006). Revenue efficiency and cost allocation. Comments submitted to the Federal Energy regulatory Commission, May 25, available at http://ksghome.harvard.edu/

Holttinen H., 2004. The impact of large scale wind power production on the Nordic electricity system, Dissertation for the Degree of Doctor of Science in Technology, VTT Technical Research Center of Finland – Helsinki University of Technology, December 2004

Klessmann C., Nabe C., Burges K., 2008. Pros and cons of exposing renewables to electricity market risks – A comparison of the market integration approaches in Germany, Spain, and the UK, Energy Policy, 36(10), October 2008, Pages 3646-3661

Littlechild S., 2007. Electricity cash-out arrangements, 9 March 2007, available at http://www.cornwallenergy.com

Newbery D., McDaniel T., 2002. Auctions and trading in energy markets: an economic analysis, Working paper University of Cambridge, October 2002

Nordel, 2007. Proposed principles for common balance management, November 2007, Available at http://www.nordel.org

Nordic Energy Regulators (NordREG), 2008. Towards harmonised Nordic balancing services – Common principles for cost allocation and settlement, Report 3/2008, Available at http://www.nordicenergyregulators.org/

Ragwitz, M., Held, A., Resch, G., Faber, T., Haas, R., Huber, C., Morthorst, P.E., Jensen, S.G., Coenraads, R., Voogt, M., Reece, G., Konstantinaviciute, I., Heyder, B., 2007. Assessment and optimisation of renewable energy support schemes in the European electricity market, Final Report, Karlsruhe.

Saguan M., Glachant J-M., 2007. An Institutional Frame to compare Alternative Market Designs in EU Electricity Balancing, Working Paper University of Cambridge, January 2007

Saguan M., 2007. L'Analyse économique des architectures de marché electrique. Application au market design du « temps réel », Thèse pour le Doctorat en Sciences Economiques, Université Paris-Sud 11, Faculté Jean-Monnet, April 2007

Tractebel Engineering and Katholieke Universiteit Leuven, on behalf of the European Commission (2009): Study on interaction and dependencies of Balancing Markets, Intraday Trade and Automatically Activated Reserves, available at http://ec.europa.eu/energy/index_en.htm

Tradewind, 2009. Integrating wind – Developing Europe's power market for the large-scale integration of wind power, Final report, February 2009, Available at http://www.trade-wind.eu/

VTT Technical Research Center of Finland, on behalf of IEA Wind Task 25, 2007. Design and operation of power systems with large amounts of wind power, State-of-the-art report, October 2007, Available at http://www.ieawind.org

Weber C., 2009. Fluctuating wind energy and adequate market design for European power systems, this Energy Policy issue

Wilmar, 2005. Distribution of the integration costs of wind power, Wilmar Deliverable 7.1, November 2005, Available at http://www.wilmar.risoe.dk

[*] Corresponding author : Tel. +3216321722 ; Fax : +3216321985 ; Email: leen.vandezande@esat.kuleuven.be

[a] MPu = marginal price of upward regulation; MP_d = marginal price of downward regulation

[b] AP_u = average price of upward regulation; AP_d = average price of downward regulation; P_{DA} = day-ahead power exchange price

[c] MP_u = marginal price of upward regulation; MP_d = marginal price of downward regulation; $component_{cap}$ = additive component

[1] However, with increasing wind power penetration, the current method for determining the amount of primary reserves – e.g. in UCTE primary reserves are currently dimensioned to handle a maximum power deviation of 3000 MW, amounting to the simultaneous outage of its two largest generation units – may be brought under discussion as another method – e.g. dimensioning according to the largest possible variation – becomes more appropriate.

[2] Note that services being specifically reserved to deal with exceptional disturbances are not further considered in this article either.

[3] BRPs facing a short position or negative imbalance in real time, took off the grid (consumption, exports and sales) more than they injected (generation, import and purchases) during a certain settlement period. BRPs with a long position or positive imbalance on the contrary injected more than they took off.

[4] In other words, assuming for instance that $MP_u = 1.3*P_{DA}$ and $MP_d = 0.8* P_{DA}$, over-contracting in the wholesale market and under-nominating of injections would also occur in a one-price system in the examples below – though to a lesser extent than in a two-price system – and this effect would be completely due to the relatively higher prices for upward regulation. This 'natural tendency' should not be reinforced through the introduction of additional penalties.

[5] Also when the BRP's bid on the balancing market is not selected in case of a negative system imbalance – implying its activation cost exceeds the marginal price for upward regulation at that moment –, the BRP will prefer to pay imbalance charges to the TSO instead of using its own services to be balanced.

[6] For instance a price cap on the real-time energy bids of secondary and tertiary reserves in Belgium

[7] It is important to distinguish this *additive* component from the *multiplicative* component as discussed previously in the context of two price systems. Contrary to the latter, the former is (a) cost-reflective, (b) not used for penalisation purposes and (c) not giving rise to undesirable BRP behaviour.

[8] Note that the discriminatory impact of an additive component on new entrants may also be counteracted by the implementation of a separate imbalance settlement system for generation and load.

[9] The arguments summarised here are similar to those referred to in discussions on the (in)efficiency of capacity markets for adequacy purposes, such as those in Finon et al. (2008).

Table 1: Imbalance settlement through a typical one-price system

		System imbalance	
		NEGATIVE (short)	POSITIVE (long)
BRP imbalance	NEGATIVE (short)	+ MP_u	+ MP_d
	POSITIVE (long)	- MP_u	- MP_d

Table 2: Imbalance settlement through a typical two-price system

		System imbalance	
		NEGATIVE (short)	POSITIVE (long)
BRP imbalance	NEGATIVE (short)	+ AP_u*(1+ $penalty_u$)	+ P_{DA}
	POSITIVE (long)	- P_{DA}	- AP_d/(1+ $penalty_d$)

Table 3: Input data for examples – One-price system

		System imbalance	
		NEGATIVE (short)	POSITIVE (long)
BRP imbalance	NEGATIVE (short)	+ 1.5*P_{DA} = **1.5**	+ 0.5*P_{DA} = **0.5**
	POSITIVE (long)	- 1.5*P_{DA} = **-1.5**	- 0.5* P_{DA} = **-0.5**

Table 4: Input data for examples – Two-price system

		System imbalance	
		NEGATIVE (short)	POSITIVE (long)
BRP imbalance	NEGATIVE (short)	+ 1.5*P_{DA}*(1+ 0.4) = **2.1**	+ P_{DA} = **1**
	POSITIVE (long)	- P_{DA} = -**1**	- 0.5* P_{DA} /(1+ 0.25) = -**0.4**

Table 5: Example on the impact of imbalance pricing on wholesale trade

	Total expected costs	
	ONE-PRICE SYSTEM	**TWO-PRICE SYSTEM**
Purchase DA = 90	**100**	**105.5**
	= 90 + 0.5*20*(1.5+0.5)/2	= 90 + 0.5*20*(2.1+1)/2
Purchase DA = 100	**100**	**104.25**
	= 100 + 0.5*10*(1.5+0.5)/2 - 0.5*10*(1.5+0.5)/2	= 100 + 0.5*10*(2.1+1)/2 - 0.5*10*(1+0.4)/2
Purchase DA = 110	**100**	**103**
	= 110 - 0.5*20*(1.5+0.5)/2	= 110 - 0.5*20*(1+0.4)/2

Table 6: Example on the impact of imbalance pricing on the provision of balancing services

	Total expected income	
	ONE-PRICE SYSTEM	**TWO-PRICE SYSTEM**
Sell 10 to TSO	**0**	**-4.25**
	= 0.5*10*(1.5-1.5) - 0.5*10*(1.5+0.5)/2 + 0.5*10*(-1.5+(-0.5))/2	= 0.5*10*(1.5-1.5) - 0.5*10*(2.1+1)/2 + 0.5*10*(1+0.4)/2
Keep 10 for own use	**0**	**-2.75**
	= -0.5*10*(1.5+0.5)/2 + 0.5*10*(1.5+0.5)/2	= -0.5*10*(1.5+1)/2 + 0.5*10*(1+0.4)/2

Table 7: Example on the impact of imbalance pricing on nominations

	Total expected income	
	ONE-PRICE SYSTEM	**TWO-PRICE SYSTEM**
Sell DA = 90	**+100**	**+97**
	= 90 + 0.5*20*(1.5+0.5)/2	= 90 + 0.5*20*(1+0.4)/2
Sell DA = 100	**+100**	**+95.75**
	= 100 - 0.5*10*(1.5+0.5)/2 + 0.5*10*(1.5+0.5)/2	= 100 - 0.5*10*(2.1+1)/2 + 0.5*10*(1+0.4)/2
Sell DA = 110	**+100**	**+94**
	= 110 - 0.5*20*(1.5+(0.5))/2	=110 - 0.5*20*(2.1+1)/2

Table 8: Allocation of capacity payments via the imbalance price

		System imbalance	
		NEGATIVE (short)	POSITIVE (long)
BRP imbalance	NEGATIVE (short)	$+ MP_u + component_{cap}$	$+ MP_d + component_{cap}$
	POSITIVE (long)	$- (MP_u - component_{cap})$	$- (MP_d - component_{cap})$