Balancing power and variable renewables: Three links

Lion Hirth a,b,c, Inka Ziegenhagen d
a Neon Neue Energieökonomik GmbH (neon), Berlin, Germany
b Mercator Research Institute on Global Commons and Climate Change (MCC), Berlin, Germany
c Potsdam Institute for Climate Impact Research (PIK), Potsdam, Germany
d Prognos AG, Berlin, Germany

ABSTRACT
Balancing power is used to quickly restore the supply-demand balance in power systems. The need for this tends to be increased by the use of variable renewable energy sources (VRE) such as wind and solar power. This paper reviews three channels through which VRE and balancing systems interact: the impact of VRE forecast errors on balancing reserve requirements; the supply of balancing services by VRE generators; and the incentives to improve forecasting provided by imbalance charges. The paper reviews the literature, provides stylized facts from German market data, and suggests policy options. Surprisingly, while German wind and solar capacity has tripled since 2008, balancing reserves have been reduced by 15%, and costs by 50%.

Keywords: Balancing power, Control power, Regulating power, Variable renewables, Wind power, Solar power, Market design

Contents
1. Introduction 1036
2. Fundamentals of balancing systems 1037
2.1. Balancing what? 1038
2.2. Roles and responsibilities 1038
2.3. Types of balancing power 1038
2.4. TSO cooperation 1038
3. Balancing reserve requirement: The impact of VRE 1039
3.1. Reserve sizing methodologies 1039
3.2. Sources of imbalances 1039
3.3. Probabilistic reserve sizing in Germany 1040
3.4. Modeling the impact of VRE on reserve requirements 1040
3.5. The German experience: A paradox? 1041
3.6. Policy options: Advancing reserve calculations 1042
4. Balancing power market: Enabling VRE participation 1042
4.1. Balancing power market design 1042

We would like to thank Catrin Jung-Draschil, Swen Löpken, Bart Stoffer, Mats Nilsson, Thorbjørn Vest Andersen, Philipp Hanemann, Susann Wöhltre, Bastian Rühle, Eckart Boege, Dania Röpke, Viktoria Neimane, Maryam Hagh Panah, Kathrin Goldammer, Benjamin Bayer, Dominik Schäuble, Hannes Peinl, Rolf Englund, Oliver Tietjen, Set Persson, Felix Buchholz, Christian Andersson, Michael Pahle, Brigitte Knopf, Fabian Joas, Falko Ueckerdt, Eva Schmid, Fredrik Carlsson, Matthias Klapper, Kristian Gustafsson, Ralf Kirsch, Sundar Venkataraman, Mike O’Connor, Aidan Tuohy, Alexander Zerrahn, Eckehard Schulze, and the anonymous referees for inspiration discussions and valuable comments. Maximilian Schumacher and Julian Bauer provided helpful research assistance. The usual disclaimer applies. The paper has been presented at the YEEES, Strommarktreften, Enerday, Euroforum, and EEM conferences. Parts of earlier versions of this article have been published as conference proceedings and in Hirth and Ziegenhagen [59,60].

* Correspondence to: Neon Neue Energieökonomik GmbH, Karl-Marx-Platz 12, 12043 Berlin, Germany. Tel.: +49 1575 5199715.
E-mail address: hirth@neon-energie.de (L. Hirth).
URL: http://www.neon-energie.de (L. Hirth).

http://dx.doi.org/10.1016/j.rser.2015.04.180
1364-0321/© 2015 Elsevier Ltd. All rights reserved.
1. Introduction

Electricity generation from variable renewable electricity sources (VRE), such as wind and solar power, has grown rapidly during recent years and is expected to continue to grow. The fact that these generators are distributed, non-synchronous, and weather-dependent causes specific challenges when integrating them into power systems [54,68,71]. With increasing amounts of VRE in many countries, system integration has become a major public policy debate with a particular emphasis on the stress that forecast errors put on balancing systems.

Balancing power is used to stabilize the active power balance of integrated power systems over short time scales from seconds to hours. In AC power systems, the demand–supply balance has to hold at every instant of time to ensure frequency stability at, usually, 50 Hz or 60 Hz. Frequency deviations have a number of problematic consequences, one being that they can mechanically destroy rotating machines such as generators. Technical procedures and economic institutions have evolved to prevent frequency instability, and the most important of these is “balancing power”.

Electricity generation from VRE has been growing rapidly in many countries, driven by technological progress, economies of scale, and deployment subsidies. Global solar PV capacity has reached 140 GW, a fourteen-fold increase since 2007, and there has been a four-fold increase in wind power to 320 GW [100]. Several power systems now accommodate VRE shares of 15% to 40%, including Eastern Inner Mongolia, Denmark, Spain, Portugal, Ireland, Lithuania, and Germany. The IEA [72] projects that medium-term growth will continue and long-term models forecast that VRE shares by 2050 will need to be several times higher than today [47,84,81]. Hence, system integration will remain challenging.

This review paper explores the interaction of wind and solar power at best; in many circumstances other factors had a larger impact. Second, the design of balancing power markets constitutes an unnecessary entry barrier to this market, and prices in balancing and imbalance markets do not regularly reflect marginal costs. With appropriate market design, VRE wind and solar not only consume but can also provide balancing services, although this requires policies and markets to be appropriately designed; balancing services supplied by VRE generators obviously reduce the pressure on the balancing system. Finally, the imbalance price which is the financial penalty for forecast errors, determines the size of forecast errors given its influence on the behavior of VRE generators. If set correctly, the imbalance price can stimulate more accurate forecasting and incentivize VRE generators to improve system stability. The three links are not independent, and policy that is only targeted at one of the channels is likely to be suboptimal. We hope our holistic view generates a more comprehensive understanding of the balancing challenge and explores the entire solution space for policy makers, market participants, and system operators.

The aim of this paper is to stimulate and structure the discussion on the interaction between VRE and balancing power. It particularly targets practitioners such as policy makers, regulators, system operators, market participants, and VRE investors. It aims to provide an overview of topics, a guide through the literature, and a summary of policy proposals. We complement the literature review with model results and empirical data where appropriate. We study German markets and policies in detail, but most observations and recommendations also apply to other European markets, and, albeit to a lesser degree, to U.S. power systems. Germany is an interesting case not only because of the extent and rapid growth of VRE penetration, but also because of data availability, cooperation between system operators, and market design reforms.

We believe that three broad conclusions can be drawn from this review. First, we find the impact of VRE on the balancing system to be less dramatic than sometimes believed. VRE growth has had moderate impact on volumes and costs of balancing power at best; in many circumstances other factors had a larger impact. Second, the design of balancing power markets constitutes an unnecessary entry barrier to this market, and prices in balancing and imbalance markets do not regularly reflect marginal costs. With appropriate market design, VRE wind and solar not only consume but can also provide balancing services. Finally, VRE and balancing systems interact via various channels. Policies interact with each other and should be considered in context. Usually, multiple policies exist to achieve the same objective, a fact that might ease implementation challenges.

The paper is organized as follows. Section 2 gives an overview of balancing systems. Sections 3–5 review the three links: reserve requirements, balancing power markets, and imbalance settlement. Section 6 concludes.

2. Fundamentals of balancing systems

We define the balancing system (or “balancing regime”) as the set of technical and economic institutions that are used to maintain...
and restore the short-term active power balance in integrated electricity systems. Other system (ancillary) services, such as reactive power compensation (voltage support) or transmission congestion management (re-dispatch), are not within the scope of this article. The balancing system comprises two economic mechanisms: the "balancing power market" to acquire balancing power, and the "imbalance settlement system" to financially clear the imbalances.

This section outlines some principles of balancing systems. We clarify the roles of different actors and present an overview of the types of balancing power used in continental Europe (UCTE), Rebours et al. [101,102], TenneT [112], ENTSO-E [36], Ela et al. [31], and Cognet and Wilkinson [20] provide international comparisons. Vandezande et al. [116] discuss economic aspects of market design. Kristiansen [83] and Bang et al. [5] provide a comprehensive survey of the Nordic balancing system while Ela et al. [31,32] and NERC [95] do the same for American systems. Consentec [25] provides, in German, a descriptive overview of the German balancing system.

This article focuses on electricity. Natural gas markets feature a similar system of balancing energy to which many of the general arguments apply [78,2].

2.1. Balancing what?

Two geographic entities are involved in balancing: the synchronous system (interconnection) and balancing areas (control areas) within the synchronous system (Fig. 2). The synchronous system is a geographic area usually equivalent to several countries, and is characterized by a common steady-state frequency. The continental United States features three synchronous systems (Eastern and Western Interconnect, ERCOT), while Europe features five synchronous systems, of which the UCTE is by far the largest.

Balancing areas are regions, usually of the size of countries, for which one system operator is responsible. Balancing systems are meant to balance both the synchronous system and each balancing area. In the language of control theory, two variables with two set
points are targeted by the balancing system: frequency, and the imbalance of each balancing area. The synchronous system is balanced if frequency is at its nominal value (50 Hz). The balancing area is balanced if net import balance is at its scheduled value (the area control error is zero).

2.2. Roles and responsibilities

In Europe, four types of actors interact in balancing systems: balance responsible parties, transmission system operators, suppliers of balancing power, and regulators.

Balance responsible parties (BRPs) or “program responsible parties” are market entities that have the responsibility of balancing a portfolio of generators and/or loads. BRPs can be utilities, sales companies, and industrial consumers. Each physical connection point is associated with one BRP. BRPs deliver binding schedules to system operators for each quarter-hour of the next day, and are financially accountable for deviations from these schedules.

Transmission system operators (TSOs) operate the transmission network and are responsible to balance injections and off-takes in their balancing area. TSOs activate balancing power to physically balance demand and supply if the sum of BRP imbalances is non-zero. Specifically, TSOs have four obligations:

1. determine the amount of capacity that needs to be reserved as balancing power ex ante
2. acquire that capacity; determine its price (capacity and/or energy) ex ante
3. activate balancing power; determine the imbalance price (energy) in real time
4. financially clear the system and allocate costs (via imbalance price and/or grid fees) ex post.

Suppliers of balancing power supply reserve capacity, and deliver energy if dispatched by the TSO. They are obliged to deliver energy under pre-specified terms, for example within a certain time frame, with certain ramp rates, and for a specific duration. Suppliers are traditionally generators, but can also be consumers. Typically, suppliers of balancing power receive a capacity payment (€/MW·h) because capacity reservation incurs opportunity costs, and/or energy payment (€/MWh) since activation is costly.

Regulators determine the balancing power market design. They also monitor market power and prescribe the pricing formula of the imbalance price. Unlike most in other markets, the rules that govern trade of balancing power are set by authorities and have not emerged bottom-up from market interaction.

2.3. Types of balancing power

Characteristics, classification, and nomenclature of balancing power vary across power systems. Since multiple sources of imbalances exist with different characteristics (see Section 3.2), several types of balancing power, or balancing power products, are employed simultaneously in most power systems. Balancing power types can be distinguished along several dimensions:

- purpose (operating/non-event v. contingency/event reserve)
- state of supplying power plant (spinning v. stand-by reserve)
- target system (synchronous system v. balancing area)
- response time (fast v. slow)
- activation frequency (direct/continuously v. scheduled)
- way of activation (manual v. automatic)
- positive, negative, or both (upward v. downward v. symmetric)

In the UCTE, balancing power is called “control power” [115], and three different types are used: primary control, secondary control, and tertiary control. They differ in purpose, response time, and the way they are activated (Table 2).

Primary control power (PC) can be fully deployed within 30 s. Being a shared resource within the UCTE, it is not activated by TSOs but by locally measured frequency deviation. At the power plant level, activation is implemented as turbine-governor control. PC is activated proportionally to the frequency deviation from 50 Hz up to full activation at deviations of 200 mHz. It is calibrated such that imbalances are replaced and frequency drops are contained (but not restored). Small deviations below 10 Hz do not lead to activation (dead band). PC can be classified as a fast, automatic, symmetric spinning reserve that is used to balance the synchronous system. It is sized as a contingency reserve (0), but also used for operational imbalances.

Secondary control power (SC) has to be available within 5 min after activation. It is activated automatically and centrally by TSOs via an automatic generation control (AGC) signal that is updated every few seconds. SC is used to restore nominal frequency, and to re-balance the respective balancing area, such that at steady-state, imbalances are replaced by SC from the same balancing area. SC can be supplied by some stand-by hydro plants, but is mostly provided by synchronized thermal generators. Hence, it is an automatic reserve with direct activation that balances both the synchronous system and the balancing area up and down; to a large extent, it is a spinning reserve.

Tertiary control power (TC), or minute reserve, is used to replace SC over time. It is either directly activated or supplied in schedules of 15 min. Activation is a decision taken by TSO staff based on current and expected deployment of SC. TC is mostly supplied by stand-by generators. UCTE [115] and Rebours et al. [101] provide more technical details on the three types of balancing power.

2.4. TSO cooperation

To integrate European power markets, the European Union aims to harmonize and integrate European balancing systems and markets. If implemented as planned, the European balancing system and all of its markets will significantly change in the coming years.

### Table 1

<table>
<thead>
<tr>
<th>Geographic scope</th>
<th>Synchronous system</th>
<th>Balancing area</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target variable</td>
<td>Entire system (continental Europe)</td>
<td>Several sub-systems (countries)</td>
</tr>
<tr>
<td>Set point</td>
<td>Frequency</td>
<td>Area control error (ACE)</td>
</tr>
<tr>
<td>Rationale</td>
<td>Avoid damages caused by frequency deviation</td>
<td>Zero</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Avoid interconnector overload; “polluter pays” principle</td>
</tr>
</tbody>
</table>
### 3. Balancing reserve requirement: The impact of VRE

This section discusses the impact of VRE on balancing reserve requirements. We give an overview of reserve sizing methods and explain which variables cause system imbalances. We then discuss the impact of VRE forecast errors on the balancing reserve, and present new model results. Although studies disagree about the size of the impact, all models show that more VRE capacity tends to increase balancing requirements. Empirically, however, the German case seems to prove theory wrong: balancing reserves could be reduced while VRE capacity increased. Finally, we suggest a number of policy options.

#### 3.1. Reserve sizing methodologies

TSOs need to determine the balancing reserve *ex ante*. The methodologies they use vary across types of balancing power and between TSOs. Holttinen et al. [69] provide an overview of methodologies. These can be distinguished by two dimensions: probabilistic v. deterministic, and static v. dynamic.

The idea behind deterministic approaches is to size the reserve according to a specific event, such as the largest credible contingency (*N–1* criterion). These approaches, however, do not account for less severe events, their probability, or correlation between sources of imbalances. The idea of probabilistic (stochastic, statistic) methods is to size the reserve such that a certain, predefined level of system reliability is met. These methods estimate the probability density function of system imbalances and use the reliability target as a cut-off to determine the size of the reserve. Probabilistic approaches require detailed knowledge of sources of imbalances, their probability distribution, and their correlation.

Reserves can be determined for long time periods such as one year (static sizing) or more frequent periods depending on the current or expected status of the system (dynamic sizing). Deterministic sizing is usually static; probabilistic sizing can be static or dynamic.

In continental Europe, reserve requirements are codified in UCTE [115]. For PC, UCTE prescribes a common European deterministic-static approach: 3000 MW are reserved, in order to compensate the loss of two large nuclear reactors connected to the same bus bar. For SC and TC, UCTE suggests a number of different approaches, but leaves the decision to the TSOs. As a consequence, the level of SC+TC reserves vary widely—from 5% of average load in France to 14% in Belgium [20]. German TSOs use a static–probabilistic approach, which we will discuss in the following.

#### 3.2. Sources of imbalances

System imbalances can stem from different sources. One way to categorize them is to distinguish stochastic from deterministic processes (Table 3). Stochastic processes are unplanned outages and forecast errors. Deterministic processes are the deviations between the stepwise (discrete) schedules and continuous

---


Variables that cause system imbalances.

<table>
<thead>
<tr>
<th></th>
<th>Stochastic</th>
<th>Deterministic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thermal and hydro generation</td>
<td>Unplanned plant outages</td>
<td>Schedule leaps</td>
</tr>
<tr>
<td>VRE generation</td>
<td>Forecast errors</td>
<td></td>
</tr>
<tr>
<td>Interconnectors</td>
<td>Unplanned line outages</td>
<td></td>
</tr>
<tr>
<td>Load</td>
<td>Forecast errors</td>
<td></td>
</tr>
</tbody>
</table>

![Image of discrete schedules causing imbalances](image1)

**Fig. 3.** Discrete schedules cause imbalances (illustration). Quarter-hourly schedules cause smaller imbalances than hourly schedules.

![Image of average German system imbalance](image2)

**Fig. 4.** Average German system imbalance for every minute of the day during the year 2011 (4 s-data from TSOs). One can identify clear patterns just before and after full hours. These schedule leaps are quite large compared to the size of reserves of about 4.5 GW. Schedule leaps can also be seen in observed grid frequency; frequency deviations are clustered around the full hours during the morning and the evening ramp (http://www.netzfrequenz.info/auswertungen/laengezeitverlauf-der-netzfrequenz.html). See also Weißbach and Welfonder [120].

The probability of unplanned outages of power plants and transmission lines (contingencies) is a function of equipment characteristics (age, type) and operational decisions (maintenance, ramp rates). Similarly, the size of forecast errors is a function of resource characteristics (such as size and distribution of wind parks) and operational decisions (forecast quality, portfolio management). Operational decisions are subject to the economic incentive provided by the imbalance system (Section 5): if imbalance prices are high, owners will operate their assets in a way that forecast errors and outages rates will be reduced.

Load forecast errors are often assumed to be normally distributed [22]. Jost et al. [74] discuss characteristics of wind power and load forecast errors. Bruninx and Delarue [13] report that wind power forecast errors can be approximated well with a Lévy α-stable distribution. Zhang et al. [121] report that wind and solar power forecast errors are weakly negatively correlated; see also Braun et al. [10].

Next to these inherently stochastic processes, there is a deterministic source of imbalances: deviations resulting from the way contracts are designed in liberalized electricity markets. Schedules are specified as discrete step functions in intervals of, typically, 15 min. However, physical demand and supply changes are smooth (Fig. 3). The differences between physical and scheduled values are called schedule leaps. These leaps apply to all types of portfolios and are substantial in size. Fig. 4 shows that schedule leaps are large and that deviations are greatest around full hours, indicating that many BRPs use hourly rather than quarter-hourly schedules (see [120,23]).

The literature often distinguishes two time scales in which imbalances occur: deviations of the dispatch interval mean from schedules (“forecast error”) and variations around the mean during the dispatch interval (“intra-dispatch interval variability” or “noise”). All four sources listed in Table 3 result in both forecast errors and noise.

### 3.3. Probabilistic reserve sizing in Germany

The German TSOs use a probabilistic approach to determine SC and TC capacities, sometimes called the “Graf/Haubrich approach” [22,23,85]. It is based on the idea of statistical convolution.

In statistical terms, the balancing area imbalance follows the joint distribution of the individual factors’ distribution functions and the reserve is set according to a predefined percentile (security level) of that function (Fig. 5).

According to the most recent publicly available document on reserve sizing in Germany [23], power plant outages are based on historical outage rates as reported by the VGB PowerTech database, schedule leaps are modeled, and forecast errors are estimated from historical system imbalances. The distribution functions of individual imbalance sources are not explicitly estimated.

### 3.4. Modeling the impact of VRE on reserve requirements

If VRE forecast errors are uncorrelated to other factors, additional wind and solar capacity *ceteris paribus* increases the size of balancing reserves. While some American studies argue that PC-type reserves would also increase [12], European studies usually find that only SC and TC reserves would be affected. An important assumption for any reserve impact study is the future improvement of generation forecasts, which are often assumed to be substantial.

UKERC [114, Holttinen et al. [68], and Brouwer et al. [12] provide surveys of the international modeling literature, being mostly studies of wind and solar integration. Holttinen et al. report that in predominately thermal power systems, most studies find that reserves increase by 2–9% of the additional wind capacity (20–90 MW per GW of wind power). Brouwer et al. report a wider range, but high estimates around 20% tend to arise in older studies. DLR [29] report 4% of additional VRE capacity in Germany, given a mix of solar and wind generation and significant improvements in forecast quality. DENA [27] report a higher value for Germany. NREL [96], a high-quality study for the Western United States, report a reserve increase of 4% of additional VRE capacity. De Vos et al. [119] assess wind power forecast errors in isolation and consequently find a much larger impact on reserves.

Ziegenhagen [122] provides a convolution-based assessment of the impact of VRE on reserve requirements. She finds that reserve requirements are increased by 6% of installed wind or solar capacity, assuming a moderate reduction of forecast errors by 30%. This number is reduced to 4% if both technologies are deployed simultaneously. Without forecast improvements, such a mixed expansion would increase reserve needs by 6.5%. If forecast errors are improved by 60%, the impact on reserves would be reduced to 1.5%. For up to 100 GW of additional capacity Ziegenhagen estimates the impact of reserve requirements to be roughly linear (Fig. 6).

While there is disagreement in the modeling literature about the size of the impact of VRE on balancing reserves, there is a consensus that additional VRE capacity increases the reserve requirement. In the following, we review empirical market data from Germany, which seems to prove this common knowledge to be wrong.
3.5. The German experience: A paradox?

Since 2008, German VRE capacity has grown from 27 GW to 78 GW. Wind and solar power now deliver 15% of consumed electricity, up from 7% in 2008. Over the same period, TSOs reduced balancing reserves by 15% (Fig. 7). This empirical fact seems to contradict common sense as well as the model results presented above.

Of course these descriptive statistics do not imply that additional wind and solar power reduce the balancing reserve requirement. What it does indicate is that other factors must have overcompensated for the VRE expansion. There are several candidates:

- improvements of wind and solar forecasts
- improvement of load forecasts
- reduced frequency of plant outages
- TSOs might have become more cost-conscious and decreased additional internal security margins
- TSO cooperation in reserve sizing
- improved intra-day market liquidity, and 15 min trading on power exchanges becoming common, allowing better portfolio management.

A quantitative assessment of these drivers is a promising direction of future research. Brouwer et al. [12] report that reserves did not increase in Denmark, Spain, and Portugal either, despite considerable VRE expansion.

During 2009/2010, the German TSOs established a balancing power cooperation (Netzregelverbund). Today, both reserve sizing and activation is done jointly such that Germany can, in practice, be treated as a single balancing area [123,124]. Since 2012, the Danish, Dutch, Swiss, Belgium, Czech, and Austrian TSOs have joined to form the “International Grid Control Cooperation” (IGCC). At this stage, the members outside Germany cooperate in terms of SC activation (imbalance netting), but size reserves individually.

In academic and policy circles, there seems to be widespread believe that wind and solar power have become major drivers for balancing power. This is reflected in the fact that there are numerous published studies that assess this relationship. Surprisingly, to the best of our knowledge, there is no ex-post estimation of the impact of the Netzregelverbund or studies that otherwise explain the decrease of the balancing reserve in Germany. More generally, the literature on the impact of the size of the balancing area on reserve requirements seems to be scarce (an exception is Milligan and Kirby [86]).

Reserve sizing is not a question of VRE alone, but of many more factors. Future studies should not assess the impact of VRE in...
isolation, but take these other factors into account. Germany’s historical experience also shows that it is possible to decrease balancing reserves while increasing VRE capacity, if system operation is organized more efficiently.

3.6. Policy options: Advancing reserve calculations

To increase the economic efficiency of reserve sizing, several proposals have been made: dynamic sizing, specific solutions for deterministic imbalances, and price-elastic procurement.

Reserve requirements could be calculated dynamically, i.e. as a function of the current or expected system conditions [69,11,79]. Dynamic sizing is currently under assessment in Switzerland [1] and Germany [74]. Bucksteeg et al. [14] report that German reserve requirements are significantly larger during the day than at night.

Empirical data shows that schedule leaks cause significant deterministic imbalances (recall Fig. 4). The purpose of balancing power is to respond to unexpected events—but schedule leaks are known to occur every day at the same time. A dedicated approach to this source of imbalance seems more appropriate. One could mandate smoother transitions between dispatch intervals, shorten intervals [120,98], or introduce a dedicated ramping product [87]. We propose to use passive balancing, i.e. to use the imbalance price to incentivize BRPs to voluntarily stabilize the system (Section 5.5).

A third possibility of improvement concerns price-elastic sizing [192]. If reserve prices are low, system operators should buy more to increase the level of reliability, and vice versa. As a side benefit, this would reduce the level of scarcity prices in energy-only wholesale markets [65].

Passive balancing and dynamic and price-elastic sizing all complement each other and all of them could be pursued. This section has discussed reserve sizing. TSOs then procure reserves on balancing power markets. We proceed by discussing these markets.

4. Balancing power market: Enabling VRE participation

Since TSOs do not own generation assets, they procure reserves on balancing power markets. Depending on the type of balancing power, suppliers receive a payment for capacity and/or energy, hence the price for balancing services is, in general, a two-part tariff. While capacity prices and energy prices for positive balancing are always non-negative (TSO pays the supplier), the energy price for negative balancing can be positive or negative (suppliers sometimes get paid for reducing output).

While the technical characteristics of different balancing power types are largely harmonized throughout the UCTE (Table 1), balancing power market design is national. A wide range of institutional setups exist, ranging from uncompensated supply obligation for generators, regulated or capped prices, mandatory offers by generators, to competitive voluntary bidding [41]. While almost all wholesale electricity prices feature marginal pricing, pay-as-bid pricing is common in balancing power markets. In contrast to wholesale electricity markets, balancing power is only demanded by the TSO, hence it is a single-buyer market.

Rebours et al. [102], ENTSO-E [36], Cognet and Wilkinson [20], and Fattler and Pellinger [42] compare market rules internationally. TenneT [112] compares the Dutch and the German market while Ela et al. [32] discuss American market design. van der Veen [117] discusses market design in the context of European Network Code development (recall Section 2.4). In the following, we will highlight relevant aspects of the German market design, report on recent market development, argue that VRE can cost-efficiently supply negative balancing, and identify barriers to entry that have prevented VRE generators from participating.

4.1. Balancing power market design

German >balancing power market design is prescribed by the regulator and has been subject to frequent changes. Shortly after the liberalization of wholesale markets in 1989, balancing power markets were created in 2001, when bilateral contracts were replaced by public auctions. Since late 2007, the four German TSOs have used a common procurement platform8. Table 4 summarizes auction design as it has been in effect since mid-2011, after the latest reform.

TSOs have a perfectly price-inelastic demand for balancing power. Bidders have to prove that they can deliver balancing power according to technical requirements (Table 1) before being allowed to participate (“pre-qualification”)10. All auctions are pay-as-bid auctions (see [89,18,93]). Bids are accepted based on their capacity price only (Leistungspreis); dispatch is done according to the energy price (Arbeitspreis). Hence, there are two independent merit-orders. PC and SC are tendered for a week, TC for each day. PC is a symmetric (bi-directional) base product, which means both upward and downward regulation has to be provided for an entire week. SC is tendered separately as positive and negative reserves for peak and off-peak periods. TC is auctioned in blocks of 4 h, again, separately for negative and positive reserves. Hence there are four SC products and 12 TC products per auction. Minimum bid sizes apply, but pooling is allowed.

The number of pre-qualified suppliers has increased significantly during the past years (Fig. 8). Today, several “unconventional” suppliers such as municipal utilities, industrial consumers, aggregators, and foreign generators are pre-qualified for all three control power types11. Bidding data is anonymous, such that the market share of new suppliers is unknown.

4.2. Market size and market development in Germany

We present empirical data of the German balancing power market in this section, for which we have compiled all balancing power bids since 200812. We discuss market size, price patterns of different types, and price trends.

Compared to the wholesale electricity market, balancing power is a niche market. Capacity payments were about €400 mn in 2014. For suppliers, this is the revenue potential of the balancing power market; for TSOs, this is the cost of capacity reservation. Divided over total electricity consumption, this corresponds to less than 0.8 €/MWh consumed. In other words, of the household price of about 300 €/MWh consumed, balancing power represents about 0.25% (Fig. 9). We estimate utilization costs to be € 200–300 million, hence capacity payments are about two thirds of the total costs for balancing; even including these costs, balancing services cost consumers little more than one €/MWh consumed. For other countries, Rebours et al. [102] report that the balancing system costs 0.5–5% of the wholesale market for electrical energy.

---

8 All data are publicly available at www.regelleistung.net, but are provided in somewhat inaccessible formats. A summary of current market developments can be found at http://neon-energie.de/Regelleistungsmarkt-2014.pdf.
consistent with the numbers reported here. Cognet and Wilkinson [20] find a similarly wide range of costs across European markets. Hence, balancing power is cheap compared to the total cost of the power system. In a sense, even if balancing is regarded as a problem for VRE deployment, economically, it is a small problem.

Table 4
Balancing power market design in Germany.

<table>
<thead>
<tr>
<th></th>
<th>Primary control</th>
<th>Secondary control</th>
<th>Tertiary control</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Auction period</strong></td>
<td>Week</td>
<td>Week</td>
<td>Day</td>
</tr>
<tr>
<td><strong>No. of products</strong></td>
<td>1 (symmetric, base)</td>
<td>4 (pos/neg; peak/off-peak)</td>
<td>12 (pos/neg; blocks of 4 h)</td>
</tr>
<tr>
<td><strong>Contract duration</strong></td>
<td>Week</td>
<td>week (peak/off-peak)</td>
<td>4 h</td>
</tr>
<tr>
<td><strong>Capacity payment</strong></td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Energy payment</strong></td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Minimum bid</strong></td>
<td>1 MW</td>
<td>5 MW</td>
<td>5 MW</td>
</tr>
<tr>
<td><strong>Number of suppliers</strong></td>
<td>21</td>
<td>28</td>
<td>41</td>
</tr>
</tbody>
</table>

Fig. 8. The number of suppliers of balancing power has grown.

Fig. 9. The cost of balancing power provision is very small when compared to the wholesale market for energy or the retail prices. Source: own calculations, Bundesnetzagentur [17].

Table 5
Correlation of prices of SC products.

<table>
<thead>
<tr>
<th></th>
<th>Peak</th>
<th>Off-peak</th>
<th>Positive</th>
<th>Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peak</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Off-peak</td>
<td>0.20</td>
<td>1</td>
<td>0.36</td>
<td>1</td>
</tr>
<tr>
<td>Positive</td>
<td>0.80</td>
<td>0.36</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>0.14</td>
<td>0.88</td>
<td>0.02</td>
<td>1</td>
</tr>
</tbody>
</table>

Fig. 10. Prices for balancing power are volatile. Demand shocks (TSO cooperation 2009), market design changes (2011) and expected low spot prices during Christmas (2013, 2014) led to price spikes.

Fig. 11. Price for negative and positive balancing (yearly average, %–change relative to 2008).

Fig. 12. Price by balancing power type (yearly average, only SC and TC, %–change relative to 2008).
PC and SC are more highly priced than TC, reflecting stricter technical requirements. As symmetric base products, PC and SC have been priced at about 20 €/MW h on average since 2008, and TC 6 €/MW·h. The fact that PC and SC are similarly priced confirms the impression that they are close technical substitutes (Table 1). Maybe surprisingly, negative balancing has been more expensive than positive balancing since 2010. The price of some products, notably positive TC, were practically zero. For both SC and TC, negative balancing tends to be more expensive in off-peak hours, while positive balancing is more expensive in peak hours. This pattern can be well explained by opportunity costs (Section 3.4). Moreover, peak prices are highly correlated with positive prices, and off-peak prices are highly correlated with negative prices (Table 5).

The prices for all three products are very volatile and price spikes occur (Fig. 10)\textsuperscript{13}. Shocks have had dramatic impacts on prices, at least in the short term. Since 2008, balancing prices have decreased significantly. Prices for TC declined more than for other products (Fig. 11) and prices for positive balancing decreased more than for negative balancing (Fig. 12).

Declining prices in conjunction with declining quantities has caused the market size to contract by half since 2008 (Fig. 13). This is bad news for suppliers, but good news for consumers. While VRE capacity has since tripled, volumes have decreased by 15%, as reported above, and costs have declined by 50% (Fig. 14). Of course, this does not indicate a causal relationship (\textit{ceteris paribus}, more VRE capacity does not decrease costs). On the contrary, it indicates that there is \textit{not} a one-to-one relationship: apparently VRE do not necessarily dominate balancing cost development, even during times of strong capacity expansion.

Explaining the evolution of prices is not trivial. Since 2008, the balancing power market has been affected by a number of shocks, which have all potentially influenced price development:

- contracted demand: this has potentially driven forecast improvements, TSO cooperation, and portfolio management which collectively outweigh the VRE capacity expansion (recall Section 3.5)
- supply shocks, such as the nuclear phase-out and the recession: these leave the wholesale market with much overcapacity that has entered balancing power markets
- lower margins on spot markets: these changed opportunity costs for thermal plants (Section 4.4)
- balancing markets becoming more competitive, possibly triggered by market design changes (recall Section 4.1; [56]) and regulatory intervention [53,57].

A more rigorous evaluation of the price development is a promising direction for further research. In the following, we indicate the drivers of the costs of thermal plants to reserve capacity for balancing, and argue that because of this cost structure, VRE would be an efficient supplier of downward balancing.

4.3. Providing balancing power with VRE generators

Wind and solar power are technically well suited for being ramped up or down very quickly, without significantly increasing maintenance costs or affecting the life-time of the plant, and with negligible IT costs [80,6,109,103,9,33,55]. In contrast, ramping of thermal plants causes temperature changes for boilers, tubes and turbines, which causes fatigue.

VRE generators can ramp down only in times when they produce electricity. To ramp up, they need to be operated below potential output, such that some generation is constantly curtailed. Consequently, while they can technically supply both positive and negative reserves, for economic reasons they are better suited to provide downward balancing.

To provide balancing capacity, one needs to use probabilistic generation forecasts to determine the share of output that is highly certain. Fraunhofer [45] has run a major joint project with German TSOs to assess the provision of balancing power by wind power. The share of wind power generation that is “firm enough” to participate in balancing power depends on the required security (probability) level and the relative size of the forecast error (without forecast errors, all generation could participate). The forecast error is a function of three major factors: (i) the size and geographic distribution of the pool of wind turbines, (ii) contract duration, and (iii) forecast horizon. A larger pool, shorter contract duration, and shorter forecast horizon reduce forecast errors. If all German wind turbines are pooled, contract duration is 1 h, and procurement occurs 1 h ahead. In more than 50% of all hours of the year wind power could supply all negative minute reserve that is required. With day-ahead auctions, this would be the case in only 10% of the hours. Görtz and Baumgart [55] apply a probability level of 99.994% and find that if all German wind power was pooled, an average of 30% of its output would be “firm”. The security level of 99.994% corresponds to the average availability of thermal plants that had bid in the balancing power market. Such forecasts are used in practice in Denmark, where wind power supplies balancing services [107].

Fraunhofer [45] also discusses the pros and cons of providing balancing power relative to a fixed schedule or relative to potential generation. Running wind turbines at fixed schedules means that some generation must be curtailed. Determining potential generation is not trivial, because of wake effects within wind parks.

\textsuperscript{13} During the Christmas week of 2013, 460 €/MW h was paid for negative SC.
4.4. The opportunity costs of reserve provision: Why VRE should participate

We will show that a thermal plant's opportunity costs of preserving balancing capacity is a function of the spot price, and that costs are especially high at very low spot prices. Hence, whenever VRE generate much electricity (allowing them to provide balancing reserves), thermal plants can only supply at high cost. This is a strong argument for encouraging VRE generators to participate in balancing power markets.

The opportunity costs of reserve provision are determined by the foregone profit from sales on the spot market. They depend on (i) the foregone status of the generator, (ii) spot market spreads, (iii) ramping and cycling costs, and (iv) part-load efficiency losses. Opportunity costs are different for positive and negative balancing. A generator that is in the money (infra-marginal) can provide negative balancing power at zero cost, because with or without balancing power provision it would be dispatched at its rated capacity. To provide positive spinning reserves, the generator has to operate constantly below its rated capacity, resulting in reduced electricity sales and part-load efficiency losses. A generator that is out of the money (extra-marginal) has to remain online despite making losses; hence its opportunity costs are avoided losses.

Ignoring ramping costs and part-load efficiency, the opportunity costs of providing positive spinning reserve, $C_t^+$, in hour $t$ can be written as a function of the spot price $p_t$, the plant's variable cost $c$, minimum load $P_{min}$ and the amount of balancing power the plant can deliver, $P_+$.

$$C_t^+ = \begin{cases} (p_t - c) & \text{if } p > c \\ (p_t - c) \cdot P_{min} / P_+ & \text{if } p < c \end{cases}$$

The opportunity costs of providing negative reserve, $C_t^-$, can be written as this:\footnote{The opportunity costs over contract duration $T$ is the sum of each hour's costs, $C_T = \sum_{t=1}^T C_t^-$: As spot prices vary, in different hours different generators have least opportunity costs. That is why longer contract duration causes inefficiencies.}

$$C_t^- = \begin{cases} 0 & \text{if } p > c \\ (p_t - c) \cdot (P_{min} + P_-) / P_- & \text{if } p < c \end{cases}$$

Figs. 15 and 16 illustrate opportunity costs of providing balancing reserves of plants with different variable costs: natural gas-fired power plants, coal-fired, and wind power.

At current European commodity prices, these plants have variable costs of around 40 €/MWh, 25 €/MWh, and zero, respectively. If the spot price equals variable costs, the power plant is at the money (marginal) and generators are indifferent to produce or not. The opportunity costs for positive balancing are zero (Fig. 15). At lower or higher spot prices, opportunity costs are positive. VRE generators have high opportunity costs because of low variable costs. The higher the foregone subsidy, the higher the opportunity costs of curtailing generation for the VRE generators. Only at very low or negative prices would it be efficient to use VRE generators to provide an upward balancing reserve.

In contrast, the opportunity costs of VRE for negative balancing (Fig. 16) are always zero. At low spot prices, the opportunity costs of thermal plants to provide balancing reserve become positive. During windy and sunny hours, VRE generation depresses the spot price, driving up the opportunity costs of thermal plants to provide negative balancing power. In other words, it is in the hours during which VRE generators are able to supply that thermal generators have high costs. In such hours it would be efficient to use VRE for downward balancing.

4.5. Policy options: Lowering entrance barriers

We have argued that VRE can efficiently supply negative balancing. However, even though the German renewable support system allows VRE to participate in balancing markets\footnote{Under the feed-in-tariff, VRE generators are not allowed to participate. Under the feed-in-premium, which now covers more than half of all capacity, they are allowed.}, wind and solar power seem not to participate in significant volumes [82]. Apart from implementation issues [108,73,45], this seems to be because the design of balancing power markets constitutes a prohibitive entry barrier. Three options exist to lower entrance barriers: adjusting the balancing power auction design; introducing a second energy auction; fostering passive balancing. Lowering entry barriers is not only important for VRE generators, but also for demand side participation [7].

The ability of VRE to provide negative balancing power is limited to times when the primary energy is available. German balancing power auctions require the provision of PC and SC for a full week. Over that time horizon, weather forecasts are too uncertain for VRE to be able to provide firm reserves. Shorter auction periods are necessary for VRE participation. We propose daily auctions with a contract duration of 1 h, similar to day-ahead spot auctions. This setup is already used in Nordic and some Eastern European countries [41]. Böttger and Bruckner [128] estimate that this would reduce PC capacity costs by 11–13% As a side benefit, it also improves the efficiency of thermal plant

$$P_{infra-marginal} = \frac{P_	ext{max}}{2}$$

$$P_{extra-marginal} = P_	ext{max}$$

$$P_{infrastructure} = \frac{P_	ext{max}}{2}$$
dispatch [75, 94, 14, 110]. It would also reduce must-run of thermal plants [61], keeping up the spot price and mitigating the market value drop of wind and solar power [58, 62]. In turn, larger quantities of renewables would be competitive [63]. The only costs involved are transaction costs, which could be reduced if the power exchange, rather than a proprietary platform, was used to procure balancing power.

In some countries, a second market for balancing energy opens after the balancing power auction is closed. In this “energy-only” market, generators are only rewarded with energy payments for providing balancing services. The gate closure of these markets is very short, such that bids can be submitted until shortly before real time. Denmark [35, 125, 107] and The Netherlands [112] are examples of such energy markets. TenneT argues that this feature is a key reason for lower costs in The Netherlands compared to Germany. It would augment the existing balancing market with an energy-only balancing market. ACER [3] proposes such a market in the “Balancing Energy” framework guideline. An alternative to energy markets is passive balancing, which we discuss in Section 5.5.

In addition, there are two proposals that would increase efficiency, but are unrelated to VRE-specific entry areas. Borggrefe and Neuhoff [8] and Sørds et al. [106] have proposed conditional bids. These are joint bids on spot and balancing power markets, for example offering balancing services only in those hours where plants are dispatched. Kahn et al. [76] and Mäggen et al. [126] argue that pay-as-bid pricing should be replaced by uniform pricing as it is more robust against uncertainty, reduces information asymmetry between companies, and mitigates market power. Several European countries use marginal pricing [41].

5. Imbalance settlement system: Setting the right incentives

Imbalance settlement occurs after the activation of balancing power (ex post). It involves two steps: the determination of the imbalance price, or “imbalance charge”, and the allocation of remaining costs or profits. The imbalance price is the price that BRP have to pay for being out-of-balance and is paid per MW h deviation from the submitted schedule (€/MW h), hence it is an energy price. This section discusses imbalance settlement, pricing rules, and the two ways in which the imbalance price can work as an economic incentive for VRE generators (and other market actors): improving forecasts, and passive balancing.

5.1. Imbalance settlement regimes

Imbalance price mechanisms are nationally regulated and differ along to differ in various aspectsewith respect to, aber nicht along several dimensions [116, 8, 34, 41]:

- two-price system (dual price) v. one-price system (single price)
- imbalance price calculated from balancing costs or from spot price
- whether or not capacity cost is included
- average v. marginal pricing
- cost-based or without punitive mark-ups (constant or variable costs; mark-ups at high imbalances, or minimal incentives)
- non-discriminatory pricing or a differentiated price for generators and loads
- settlement intervals: 15 min, 30 min, 60 min
- publication lag: time between end of settlement interval and publication of the imbalance price
- whether or not there is a legal obligation to be balanced.

The German imbalance pricing mechanism is determined by the regulator (Bundesnetzagentur) and has been adjusted several times in recent years. Since May 2010, there has been a common German imbalance price (Ausgleichsenergiepreis, reBAP). The price is determined for settlement intervals of 15 min as the average dispatch cost (net energy payments divided by net balancing energy). The system is designed to be cost-neutral in the sense that all dispatch costs are borne by unbalanced BRPs. As the energy payment in the balancing power market is subject to pay-as-bid pricing, the imbalance price is generally different from the energy price in the balancing market; in general, it does not represent the marginal costs of activating balancing reserves.

Germany uses a non-discriminatory one-price system, hence short (undersupplied) and long (oversupplied) BRPs are settled with the same price. Apparently, however, the German regulator perceived the imbalance spread as being too low to provide a sufficiently strong incentive to deter imbalances [15, 24]. Consequently, a punitive mark-up was introduced in late 2012.[16]

TSOs publish imbalance prices with a delay of several months. BRPs trade imbalances in the so-called “day after” market. In a one-price system this does not affect expected costs—but it reduces uncertainty. The costs for capacity reservation are socialized via grid fees, as in most European countries [41].

The relevant economic incentive for decisions that concern imbalances is the imbalance spread, being the difference between imbalance price and the corresponding day-ahead price. Usually, BRPs on the “wrong” side, i.e. those increasing (aggravating) the system imbalance, pay the imbalance spread, while BRPs that are on the “right” side earn a spread.

5.2. Estimates of VRE balancing costs

Numerous studies estimate the balancing costs of wind and solar power, i.e. the €/MW h cost for deviations from schedules. Holttinen et al. [68] and Hirth et al. [64] survey the literature on wind power, and Hirth [62] on solar power.

Model-based studies report costs to be between zero and 6 €/MW h, even at very high penetration rates of up to 40%. A linear trend of these estimates shows a moderate increase in balancing costs with wind penetration [54, 111, 51, 105, 26, 113, 50, 68, 88, 48].

Studies that estimate balancing costs from observed imbalance prices show a very different picture: cost estimates vary between zero and 13 €/MW h and seem to be uncorrelated with penetration [66, 99, 97, 77, 30, 64] (Fig. 17).

Peculiarities of balancing power markets and imbalance pricing might explain this apparent paradox. Market concentration can be high in balancing power markets and price variation can be large (Section 4). For example, balancing costs were 11 €/MW h in Austria in 2014, several times higher than in 2011, while costs remained stable in neighboring Germany [30]. Imbalance prices sometimes include punitive mark-ups that are unrelated to underlying costs (Section 5). This is one reason for the Dutch balancing cost of 9 €/MW h, as Pinson et al. [99] explain. One should be careful with generalizing cost estimates from tight markets or prices that do not reflect costs.

5.3. German imbalance prices

Fig. 18 displays the average imbalance spread for the years 2011–2014 as a function of the system imbalance (see also Table 6). The positive correlation between the system imbalance and the imbalance price indicates that, overall, the pricing mechanism provides an economic incentive in the right direction. When the

system was long, long BRPs lost on average 50 €/MW h, since they had paid 40 €/MW h on the day-ahead market, and received −10 €/MW h as an imbalance price. When the system was short, the imbalance spread was 52 €/MW h. In less than one percent of all quarter-hours the imbalance spread provided a perverse incentive to BRP, being negative in times of system undersupply or positive in times of system oversupply.

Surprisingly however, the imbalance price was only 35 €/MW h on average, while the day-ahead spot market price was 41 €/MW h. Hence, during these years, it would have been profitable – albeit unlawful – for a BRP to be constantly short. In other words, the imbalance market and the day-ahead market were not free of arbitrage opportunities. In one-price systems, imbalances are only costly if correlated with the system imbalance: an imbalance that is uncorrelated with the system imbalance pays an average imbalance spread of zero\(^{17}\).

Over the course of the day, the imbalance spread is everything but randomly distributed (Figs. 19 and 20). The first and last quarter-hour of each hour features significant deviation from the mean, especially during the morning and the evening ramps. For example, the average imbalance spread just before 7 a.m. is 32 €/MW h, and the average spread just after 7 a.m. is −27 €/MW h. This pattern is rooted in schedule leaps (recall Fig. 4). We will argue below that BRPs should be allowed to respond to these price incentives and balance them passively.

### 5.4. The balancing price: an incentive for better forecasting

TSOs and regulators often view the imbalance price primarily from a cost allocation perspective, i.e. as a mechanism to recover the cost of utilizing balancing reserves. However, from an efficiency perspective, the crucial role of the imbalance price as an economic incentive to BRPs is to avoid (or not avoid) imbalances. BRPs can reduce imbalances in many ways: by improving forecast tools, updating forecasts more frequently, shifting from hourly to 15 min scheduling, trading more actively on intra-day markets, and dispatching assets more accurately. Rational BRPs invest in such imbalance management measures up to the point where the marginal cost of reducing imbalances equals the

---

**Table 6**

<table>
<thead>
<tr>
<th></th>
<th>Average.</th>
<th>System long (80% of all hours)</th>
<th>System short (40% of all hours)</th>
<th>System very long (&lt; −2000 MW) (4% of all hours)</th>
<th>System very short (&gt; 2000 MW) (2% of all hours)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Imbalance price(^a)</td>
<td>35 €/MW h</td>
<td>−10 €/MW h</td>
<td>94 €/MW h</td>
<td>−43 €/MW h</td>
<td>183 €/MW h</td>
</tr>
<tr>
<td>Day-ahead price(^a)</td>
<td>41 €/MW h</td>
<td>40 €/MW h</td>
<td>42 €/MW h</td>
<td>39 €/MW h</td>
<td>50 €/MW h</td>
</tr>
<tr>
<td>Imbalance spread(^a)</td>
<td>7 €/MW h</td>
<td>−50 €/MW h</td>
<td>52 €/MW h</td>
<td>−82 €/MW h</td>
<td>132 €/MW h</td>
</tr>
</tbody>
</table>

\(^a\) Time-weighted average. Deviations caused by rounding.
marginal benefit of doing so, i.e. the imbalance spread. For statically and dynamically efficient resource allocation, the imbalance price should reflect the marginal economic costs of solving imbalances by means of balancing power. This is currently not the case.

Efficient resource allocation requires the imbalance price to represent marginal, not average, costs of balancing. The combination of pay-as-bid auctions on the balancing power market and average pricing on the imbalance market leads to inefficiently low imbalance prices. Hence, either pay-as-bid payment should be replaced by marginal pricing (for more reasons to do this, see 0), or the imbalance price should be based on marginal deployment costs.

Similarly, capacity costs should not be socialized, but borne by those BRPs that caused the need for reservation [116]. Given the practical difficulties of implementing such a pricing rule under uncertainty, a pragmatic approach could be to allocate these costs as a mark-up on the imbalance price. This would increase the imbalance spread by about 20 €/MW h in absolute terms. If reserves are sized to cover the loss of the largest unit, as is the case for PC, one might consider allocating capacity costs to this unit.

Setting the correct incentive might be more relevant for VRE generators than for other BRPs. Relative to their output, imbalances are larger for these than for other generators. Forecasting difficulties of implementing such a pricing rule under uncertainty, a pragmatic approach could be to allocate these costs as a mark-up on the imbalance price. This would increase the imbalance spread by about 20 €/MW h [116] in absolute terms. If reserves are sized to cover the loss of the largest unit, as is the case for PC, one might consider allocating capacity costs to this unit.

5.5. Passive and active balancing

When TSOs deploy balancing power, they actively balance the system. The price paid for this service is the capacity and the energy payment for balancing power. The imbalance price provides the incentive to BRPs to “passively” balance the system by purposely deviating from the schedule (“self-balancing”). TSOs can either actively balance the system by sending a dispatch signal to the suppliers of balancing power, or passively balance the system by sending a price signal to BRPs.

There are two preconditions for effective passive balancing: a timely publication of the imbalance price, and the legal ability for BRPs to respond to the price signal. Traditionally, the Dutch TSO has used this mechanism quite heavily [112], while the German TSOs have followed a philosophy of active balancing. In fact, in Germany it is illegal for BRPs to be deliberately unbalanced—even if this stabilizes the system. The German regulator and TSOs expect BRPs to stick to their schedules, not to respond to price signals.

Passive balancing is a close substitute for active balancing, especially for slow reserves. Moreover, deterministic imbalances—such as schedule leaps—could be efficiently targeted by passive balancing. Fostering passive balancing could be an alternative (indeed, a very good substitute) to the introduction of energy-only balancing markets.

5.6. Policy options: The imbalance charge as price signal

There are three major sources of inefficiency in the German imbalance market: practical and legal barriers to passive balancing, and an inefficiently low imbalance spread.

Passive balancing should be encouraged. First, it needs to be legalized, and second, the imbalance price needs to be published shortly after real time. In France, Benelux, and the UK, prices are published within 1 h [41].

Imbalance prices should reflect the marginal costs of balancing. To do so, they should include the costs of holding reserves, and they should reflect the marginal cost of deploying them. Marginal pricing could be implemented by either switching to marginal pricing on balancing markets, or by adjusting the imbalance pricing formula accordingly. Average pricing and socializing capacity cost both tend to depress the imbalance spread. In other words, the German imbalance spread is currently inefficiently low. Current market rules constitute a positive externality: unbalanced actors do not bear the full costs of their action. The incentive that BRPs receive to keep their portfolio balanced is too weak. This implies that the incentive for VRE generators to improve forecasts is too weak.

6. Concluding remarks

This paper has discussed three interfaces between variable renewables and the balancing system: reserve requirements, participation of VRE generators in balancing markets, and the incentives provided by the imbalance price. These links interact and need to be considered jointly to explore the entire solution space of policy makers. Take the example of passive balancing: allowing BRPs to respond to price signals might be a good substitute for reforms of the balancing market, and could reduce the need for holding reserves.

We believe we can draw four broad findings from this study. First, the balancing reserve requirement depends on a multitude of factors; the wind and solar power forecast error is only one of

<table>
<thead>
<tr>
<th>Table 7</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Policy options.</strong></td>
</tr>
<tr>
<td><strong>Proposal</strong></td>
</tr>
<tr>
<td><strong>Reserve requirement</strong> (Section 3)</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Balancing power market</strong> (Section 4)</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Imbalance settlement</strong> (Section 5)</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

---

17 A simulation of hundred normally distributed imbalances resulted in an average imbalance spread of 0.01 €/MW h.

18 In 2011, the costs for positive and negative capacity reservation (excluding PC) were € 160 million and € 310 millions, respectively. The amount of energy activated was 77 TW h and 18 TW h [116]. Allocating capacity costs through imbalance prices would have increased the imbalance spread by about 20 €/MW h, both in periods of undersupply and oversupply.

19 For solar power forecasting see also IEA Task 46, http://task46.iea-shc.org/.
several important drivers. Second, while German VRE capacity has tripled since 2008, reserves have been reduced by 15%. This indicates that other factors can be quantitatively more important than VRE, even in periods of strong capacity expansion. One can interpret this as an indication that balancing power is not necessarily a major barrier to VRE integration at moderate penetration rates. Third, the design of balancing power markets determines the incentives for VRE generators to provide balancing power themselves. Current market design constitutes a major barrier for participation. Finally, the design of imbalance settlement systems determines the incentives for BRPs to balance their portfolios. Specifically, it sets the incentives for VRE generators to produce accurate forecasts. Currently, the incentives for accurate forecasting are inefficiently low.

Throughout the paper we have also suggested a number of policy options, summarized in Table 7. We propose ways to improve forecasting accuracy. Unfortunately, incentives for accurate forecasting are still insufficient to produce accurate forecasts. Currently, the incentives for accurate forecasting are inefficiently low.

References


