UNIVERSITY OF CAMBRIDGE | Energy Policy Research Group

High renewable penetration: a new "tragedy of the commons"[1]

EPRG Working Paper 2036

Cambridge Working Paper in Economics 20119

David Newbery

Abstract

At high levels of wind penetration, surplus wind that cannot be exported must be curtailed. Marginal curtailment is 3-4+ times the average curtailment, but even in an efficiently designed market, price signals for wind investment are given by *average* not *marginal curtailment*, creating a "tragedy of the commons" that requires a corrective charge to restore efficiency. The paper sets out a model calibrated to Ireland in 2026, showing the source of distortion, and derives new formulae for the capacity credit of wind, the learning subsidy and corrective charge needed to deliver the efficient level of renewables penetration, and estimates of their magnitude.

Keywords wind curtailment, market failures, corrective charges

JEL Classification D41,D61, H23; H41, Q28

Contact dmgn@cam.ac.uk

Publication April 2022

[1] This paper is a shorter and simpler version of "Club goods and a tragedy of the commons: the *Clean Energy Package* and wind curtailment" (EPRG 2036) dated 31 December 2020.

High renewable electricity penetration: a new "tragedy of the commons"

David M Newbery[*†]
EPRG Cambridge

April 10, 2022

Abstract

At high levels of wind penetration, surplus wind that cannot be exported must be curtailed. Marginal curtailment is 3-4+ times the average curtailment, but even in an efficiently designed market, price signals for wind investment are given by average, not marginal curtailment, creating a "tragedy of the commons" that requires a corrective charge to restore efficiency. The paper sets out the simplest model to demonstrate this new form of market failure, showing the source of distortion. Auctioning suitable contracts solves the problem.

Keywords: wind curtailment, market failures, corrective charges

JEL Classification D41, D61, H23; H41; Q28

1 Introduction

The first theorem of welfare economics states that with a full set of markets a competitive equilibrium is Pareto efficient provided consumers are non-satiated. A competitive equilibrium exists provided production sets are convex, including the boundary case of constant returns to scale. In particular, in a stationary constant returns world of risk-neutral price-taking agents (or adequate hedging opportunities) if we can find a competitive equilibrium it should be efficient (provided all externalities and public goods are properly priced).

The electricity industry seems an ideal model of a potentially competitive market. The service of electricity supply is homogenous (as the System Operator is charged to maintain quality of service within very tight limits). In mature liberalized markets like the U.S. Standard Market

[*]Faculty of Economics, Sidgwick Ave, Cambridge, CB3 9DE, UK, ph: +44 1223 335248; email: dmgn@cam.ac.uk.

[†]I am indebted to very helpful comments from Toke Aidt, Par Holmberg, Robert Ritz, Richard Green, Stan Zachary and Chris Dent with the usual disclaimer.

Design and the EU Integrated Electricity Market producers (generators) offer quantities of goods at each possible price (their supply function) while consumers (retailers or large customers) offer their demand functions into auctions (day-ahead for each hour or half-hour, intra-day closer to delivery, and into the balancing market for real-time dispatch). The auctions clear to determine prices at each node (in the U.S.) or zone (in the EU) for each time period, simulating the Walrasian auctioneer in a way unseen in almost all other markets.

The main difference from the classic Walrasian case is that most buyers (retailers) buy on behalf of consumers, most of whom do not face a real-time price and so cannot express their instantaneous demand functions. Instead of a classic solution where the price adjusts to reduce excess demand to zero, the System Operator ensures reliable matching of supply and demand up to the "Value of Lost Load", a proxy of the maximum willingness to pay for uninterrupted delivery. Provided this is an acceptable value,[1] and provided externalities (pollution, greenhouse gas emissions) are properly priced, and provided all suppliers take prices as given, then for an electricity industry in long-run equilibrium with a sufficient fraction of controllable generation it is a standard result that the equilibrium will be efficient. This is demonstrated below in a simple model where we assume a stationary equilibrium with either risk-neutral agents or adequate hedging.

The electricity market model of this article has all the obvious externalities properly internalized. It has a socially optimal carbon price, all other damaging externalities are properly priced or prohibited, and, critically, renewable generation is assumed competitive against conventional (fossil) generation, as is increasingly the case in many countries. Any contracts for hedging risk ensure that generators face efficient market prices.[2] Variable renewable electricity (VRE) like wind and solar PV have peak to average output ratios of 3-4:1 (wind) or 4-10:1 (PV). For concreteness from now on the relevant VRE is wind and PV will be ignored (but the same analysis applies to that case too). Above a certain level of penetration, wind will inevitably be in surplus some of the time, as it will be excessively expensive to store or export all the surplus. Most current renewables support systems only pay the contracted price on delivery, encouraging renewables to offer to generate even if the price falls to zero. In such cases the System Operator must curtail some fraction of total wind to balance supply and demand. In the market studied here, when wind is surplus the market price falls to its avoidable cost, for simplicity taken as

[1] It is possible to imagine a world in which each individual has programmed all appliances to disconnect all or some level of demand at a critical price which is revealed to the supplier who then constructs a more accurate demand function. Electric vehicles and some electrical heat pumps can already accept smart, price-sensitive demands.

[2] The yardstick contract-for difference of Newbery (2021b) provides appropriate risk hedging while allowing the market price to guide output decisions.

zero. When the price falls to the avoidable cost all surplus wind will voluntarily self-curtail (i.e. reduce output) until supply and demand are matched, at which price all wind will earn zero profits.

The surprising and novel claim in this article is that even if all externalities are properly priced and the industry is in a free-entry long-run price-taking equilibrium, once the share of VRE (wind) reaches a critical level at which (self-)curtailment occurs, the resulting competitive equilibrium will be inefficient without additional taxes or charges.[3] The reason is that, as shown below, the marginal curtailment is a multiple (3-4+) of average curtailment. Consequently, free entry will result in too much VRE, as all wind turbines will equally make profits in non-curtailed hours, and all make zero profit in curtailed hours.[4] As entry will be driven by average, not marginal profit, there is a tragedy of the commons, akin to free entry of grazing cattle on the common pasture (Hardin, 1968). In this case the pasture is the number of profitable hours of operation. As there is a considerable difference in the avoidable cost of non-VRE and VRE generation, the move from uncurtailed to curtailed periods represents a step change in efficient prices, not a marginal change that is more typical of conventional markets. However, this is not the main reason, as there is no inefficiency with nuclear power with zero short-run avoidable costs. Efficient volumes of VRE therefore requires properly charging for access to these high value hours through some form of entry levy.

Nor is the problem the fundamental non-convexity of Starrett (1972). He notes that public goods can be brought within the Arrovian competitive framework by creating a set of artificial markets with individual prices (one for each recipient of the public good). For public bads like pollution and with polluters' rights, firms harmed would, at some level of pollution, avoid making negative profits by closing down, introducing a (fundamental) non-convexity in the production sets. In our example, all wind farms are identical, and the problem is not that some go out of business (or, as is the case, should not enter) but they all suffer as a result of excess entry. While this may look like a pecuniary externality, and hence completely compatible with efficiency, the result is inefficiency, or a fundamental market failure.

This article develops the simplest model to demonstrate this generic problem, and sidesteps most of the complexities of real electricity markets (set out in Newbery, 2020 with the same result under more realistic assumptions). That paper also reports the numerical estimates of the size of this entry levy based on a calibrated model of the Single Electricity Market of the island of Ireland to demonstrate that it is non-trivial (10-20% of the capital cost of VRE). The

[3] Stahn and Tomin (2021) note that the over-use of the common pool resource of artesian aquifers can be corrected by a stock-related tax.

[4] This is quite different from the excess entry of gas-fired generation caused by the entry of firms with incomplete information (Hill, 2021).

fuller background paper deals more fully with annual variability of VRE and a quantification of any learning externalities created by VRE that will be largely ignored here. The next section lays out some of the basic facts about electricity markets, the reasons why a massive expansion of VRE is to be expected in the coming decade, and the characteristics of VRE that give rise to the problem. The literature survey is confined to the problem addressed and the resulting model. A fuller survey of how electricity markets could and do address problems of reliability, VRE support and pricing, and general electricity market design is provided in Newbery (2020, 2021a).

2 Relevant features of liberalized electricity market with variable renewable electricity

While price-taking behaviour may be problematic in some liberalized electricity markets, there is no obvious reason why this could not be the case in larger markets. The European Integrated Electricity Market in 2018 had over 1,000,000 MW of installed capacity.[5] The typical sizes of generating units ranges from 5-50 MW (for peaking units) to 500 -700 MW for baseload units. Recent nuclear plants are somewhat larger, renewable generation units much smaller. Even if all units were 500 MW there would be 2,000 potentially competitive units, while allowing for different sizes by technology the number could be up to 20,000 (as 43% of 2018 installed capacity was renewable electricity with small average unit sizes). Regional markets in the US (e.g. PJM) can be about 180,000 MW and so can support a high number of competitive units. Great Britain, illustrated below, has about 100,000 MW, roughly half of which is conventional and half renewable.

Ambitious plans to decarbonize electricity will require very high levels of VRE generation, mainly on- and off-shore wind and solar PV, in many cases more than doubling existing capacity by 2030. Both decarbonization and supporting VRE are global public goods, VRE through its learning externalities that lower the cost of future investment. To solve the problem of financing such public goods, the EU requires (in its *Clean Energy Package*) member states to agree targets for emissions reduction and VRE penetration – an excellent example of turning these into club goods (Buchanan, 1965). The UNFCCC Paris Agreement and *Mission Innovation*[6] are examples of widening the club, ideally to the whole world. The EU's targets are set out in the *2030 Climate and Energy Framework*.[7] For these to be delivered in liberalized electricity markets, a number

[5] https://www.iea.org/data-and-statistics/charts/installed-capacity-in-the-european-union-2000-2010-and-projections-up-to-2040-in-the-stated-policies-scenario

[6] see http://mission-innovation.net/

[7] at https://ec.europa.eu/clima/policies/strategies/2030_en

of market failures and distortions will have to be addressed. Carbon pricing in the EU and UK has already reached Paris-compliant levels, and in early 2022 forward prices were nearly US$90/tonne. As noted, on-shore wind and PV in many member states are cheaper than new investment in fossil plant.

3 Literature Review

We are interested in high levels of VRE penetration that lead to the need for system-wide curtailment, rather than local congestion management. Heptonstall and Gross (2020) find that their comprehensive and recent review "revealed only limited data sources for aggregated costs at high VRE penetrations, with the ranges determined by assumptions made in these studies about sources of flexibility." Most studies of the impact of VRE concentrate on their price impact – the merit order effect in which low variable cost renewables push out the supply curve and lower prices. The static merit order impact of renewables capacity in displacing fossil plant is well-understood (Clò et al., 2015; Cludius et al., 2014; Deane et al., 2017; Green and Vasilakos, 2012; Ketterer, 2014; Csereklyei et al., 2019). The long-run equilibrium effect is more nuanced, depending on entry and exit decisions of conventional plant. Green and Léautier (2015) provide the most sophisticated analytical model, and this article only considers long-run equilibrium.

High VRE penetration raises particular problems for measuring their contribution to capacity adequacy and measuring their equivalent firm capacity (EFC) – the amount by which 1 MW of the considered technology can displace firm capacity (guaranteed to be present when needed) and maintain the same reliability standard. Joskow and Tirole (2007) set out the stringent conditions under which well-designed markets could deliver the specified level of reliability in markets with price caps and capacity obligations, and a mixture of price-responsive customers who can respond to real-time scarcity prices and unresponsive customers who face fixed prices. Working back from a derivation of the value of lost load (which they point out is unlikely to be independent of nature of the load-shedding event), they show in their benchmark case that all generators and Load Serving Entities should face the value of lost load in cases of load shedding. They conclude that the unusual physical characteristics of electricity and networks "makes achieving an efficient allocation of resources with competitive wholesale and retail market mechanisms a very challenging task." (Joskow and Tirole, 2007, p83).

Bothwell and Hobbs (2017, p174) argue that "many nontraditional resources have limitations that are not directly translatable into equivalent forced outage rates in adequacy calculations." They also note that "the marginal contribution of wind and solar often decreases as the installed amount increases (Keane et al. 2011)." Part of the reason is curtailment, discussed below, but a more important reason is that while failures of conventional plant are uncorrelated, wind and solar PV outputs are typically quite highly correlated with similar plant in the same region. Keane et al. (2011) is particularly relevant in underlining that the EFC of wind not only depends on the amount of wind capacity, but on the strength of the wind in any year, illustrating this for Ireland between 1999 and 2008. This dependency and its implication for the measurement of EFC has been brought more up to date in Zachary et al. (2019). That article also provides a useful discussion of the relationship between two different reliability metrics, the Loss of Load Expectation (LoLE, number of hours on average per year when load may be shed) and Expected Energy Unserved (EEU), the fraction of MWhs per year that may be shed. For many but not all purposes there is a direct mapping between them, justifying the choice of LoLE as a suitable metric (but not for the evaluation of storage). They note that VRE can be treated in the same way as conventional plant only if "the process of variable generation is statistically independent of that of demand, in which case the de-rated level of variable generation is close to its mean value" – a condition that is not satisfied in the case of high wind penetration in many countries and studied further in Newbery (2021b). The model considered here does assume that wind is uncorrelated with demand and confirms for this special case that the EFC is close to its mean value, but this should not be considered generally true.

The literature on curtailment concentrates on either local curtailment and congestion management, discussed by Joos and Staffell (2018) for Britain and Germany, or the need for storage (Pudjianto et al., 2014; Weiss and Wänn, 2013). Bothwell and Hobbs (2017) point to the potential distorting interactions between VRE support design and curtailment, and also its role in delivering reliability, partly explaining why the EFC of VRE declines with increasing penetration. At past rather low levels of penetration, Heptonstall and Gross (2020) find that "the median values for the share of VRE output curtailed across all penetration levels is consistently low, not exceeding 5%" but as this article shows, because the marginal curtailment is many times the average level, this can rapidly rise without a very flexible system. The SEM, where curtailment is already above 8%, therefore provides a foretaste of the future. In the most fully articulated dynamic model of VRE, Green and Léautier (2015) examine the marginal value of renewable capacity but only in so far as it displaces conventional generation, drives down future VRE capital costs and increases the distorting effects of the tax on energy to recover the subsidies. To the best of our knowledge there are no studies on the implications of the difference between marginal and average curtailment for market distortions.

To summarize, as far as can be determined, there are no articles that identify the market failure identified in this article. To repeat, the market failure is surprising, as it occurs when the obvious conditions for competitive equilibrium to be efficient are satisfied – all externalities internalized and convex production sets.

4 The model

The model is the simplest version to illustrate the problem. A more realistic model is available at Newbery (2020) but the results are essentially the same. In this version all obvious market failures are assumed away or internalized, so that VRE faces market prices when deciding whether or not to generate, the carbon price (relevant for fossil generation) is at the efficient level and all learning spill-overs are either zero or properly remunerated. The electricity market is isolated and has determined the Value of Lost Load (VoLL), V, which sets the reliability standard of L hours Loss of Load Expectation (LoLE) per year. To show that the problem arises because of variability and not because prices can fall to zero, initially there are three conventional (i.e. controllable or dispatchable) types of generator: peaking plant (e.g. open-cycle gas turbines), conventional fossil base-load plant (e.g. combined cycle gas turbines), and nuclear plant. In even modest-sized regions, most individual generating units are small relative to total demand, so can be considered smoothly expansible at constant cost, giving constant returns to scale. Their Equivalent Firm (or de-rated) Capacities (EFCs)[8] are P, F, and N, with annual fixed costs (to recover the capital and other fixed costs) r_J and variable operating costs v_J ($J = P$, F, and N for the plant types), with $v_P > v_F > v_N = 0$, and $r_N > r_F > r_P$.

Let $D(t)$ be demand in hour t with the Load Duration curve $D(h)$ with $D' < 0$, so that load is re-ordered with the highest load in hour 0, where h is the number of hours that demand is higher than $D(h)$. Then $D(L)$ is the required firm (de-rated) capacity required to meet the reliability standard. The market is in long-run equilibrium with all future costs and $D(h)$ known and constant. The ability of all plant to generate up to its EFC is assumed independent of its output in the preceding hour, so that in the ordered hours, h, the sole determinant of plant outputs are $D(h)$ and their relative cost. The total cost of meeting demand (except for the L hours of lost load) and the cost of the amount of lost load (in MWh), valued at the VoLL, V, is

$$\begin{aligned} C &= Nr_N + Fr_F + Pr_P + V\int_0^L (D(h) - D(L))dh + v_P PL + v_P \int_L^{h_P} (D(h) - N - F)dh \\ &\quad + v_F F h_p + v_F \int_{h_P}^{h_F} (D(h) - N)dh + v_N N h_F + v_N \int_{h_F}^{H} D(h)dh, \end{aligned} \tag{1}$$

(where the last two terms can be ignored as $v_N = 0$). Here h_p is the number of hours peak generators run, h_F is the number of hours fossil baseload plant runs, and H is the number of hours in the year (8,760). Peaking generation only runs when $D(h) \geq N + F$ and fossil generation will run all the hours for which $D(h) \geq N$. The first integral is the cost of lost load, the following

[8] Conventional plant is de-rated to allow for the independent risk of failure, so that 1 MW of EFC of any plant is equally capable of delivering the required reliability standard. National Grid (2014) explains the concept and resulting de-rating factors.

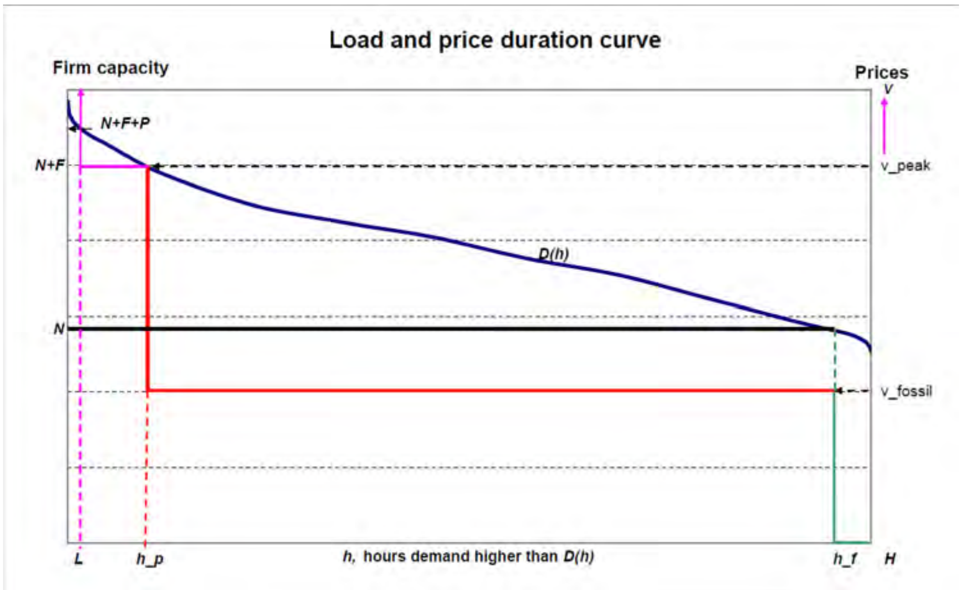

Figure 1: Load and efficient price duration curves

term ($v_P PL$) is the cost of running peaking plant for the hours that it runs at full capacity, followed by the second integral when it runs at less than full capacity, and similarly for the other technologies. Figure 1 shows the load duration curve and the amounts of firm capacity and the aligned efficient price duration curve on the right hand axis (truncated as $V >> v_P$).

The LoLE, L, satisfies $D(L) = N + F + P$, fixed by the VoLL, V in (3) below. It follows that $\partial P/\partial F = \partial P/\partial N = -1$, allowing us to ignore the choice of P in the optimization. The other two critical hours, h_P and h_F, will depend on the EFCs of the various plant types, with $D(h_P) = N + F$ and $D(h_F) = N$. The first order conditions for cost minimization of (1) are

$$\begin{aligned} 0 &= \frac{\partial C}{\partial F} = (r_F - r_P) - v_P(L + h_P - L) + v_F h_P, \\ 0 &= \frac{\partial C}{\partial N} = (r_N - r_P) - v_P h_P - v_F(h_F - h_P) + v_N h_F \\ &= (r_N - r_P) - h_F v_F - h_P(v_P - v_F), \\ h_P &= \frac{r_F - r_P}{v_P - v_F}, \quad h_F = \frac{r_N - r_F}{v_F}, \end{aligned} \tag{2}$$

which is a well-established result from screening curve analysis (e.g. see Stoft, 2002).[9] Once the critical hours are fixed, the required efficient EFCs for different plant can be deduced from $N = D^{-1}(h_F)$, $F = D^{-1}(h_P) - N$, $P = D(L) - N - F$. While the critical hours h_J depend only

[9]Screening curves are the plot of total cost of any technology against hours run. The least-cost plant mix is their lower envelope, so that peak capacity has the lowest intercept but the steepest slope and meets the base-load fossil where $r_P + h_P v_P = r_F + h_P v_F$. Similarly the fossil total cost line meets the nuclear total cost line where $r_F + h_F v_F = r_N$.

on costs, the capacities and amount generated by each type of plant depend on demand.

4.1 Decentralizing the efficient solution

Efficient prices when there is adequate capacity to meet demand, p_J, will be equal to the short-run system marginal cost (SMC) set by the most expensive plant required to meet demand. In this constant returns case $p_J = v_J$, $J = P, F, N$. When load is shed to balance demand and supply (in the L shortage hours) the price will be set by the demand side at the Value of Lost Load, (VoLL), $p_L = V$. Peaking generation only runs for h_P hours and must cover its cost. It only makes profits when prices are higher than its avoidable cost, which, except for lost load hours, will also be the price. This leads to a very simple (and widely recognized) relationship between the VoLL, V, the LoLE, L, and the net cost of new entry (net CoNE, i.e. net of any revenues earned in the market):

$$r_p = L(p_P - v_p) + (h_P - L)(p_P - v_p) = L(V - v_P). \tag{3}$$

The net unit profit of base-load fossil plant will be, after cancelling zero terms where $p_F = v_F$ and substituting for V from (3) and h_P from (2):

$$\begin{aligned}\Pi_F &= L(V - v_F) + (h_p - L)(v_P - v_F) - r_F, \\ &= r_p + h_p(v_P - v_F) - r_F = 0.\end{aligned}$$

In other words, free entry that drives net profit to zero delivers the efficient volume of this capacity. The same is readily shown to be true of nuclear net profits (where free entry is interpreted as "subsidy-free" contracted entry).

5 High penetration of Variable Renewable Electricity

In many markets (e.g. Queensland, see Simshauser, 2022) the cost of solar PV has fallen so far that it is very competitive against fossil generation, and households are willing to install solar panels without subsidy. Similarly on-shore wind and increasingly off-shore wind are becoming competitive. As such they often seek "subsidy-free" long-term contracts.[10] Both technologies are variable, or intermittent, and cannot be dispatched above the varying (insolation or wind speed) resource-limited output. Climate change plans expect a high level of Variable Renewable Electricity (VRE) in the next decade, and as peak output is typically 2-4 times average output (for wind) and 3-10 times average (for PV), peak output will exceed demand for some hours once

[10] In this perfect foresight stationary world long-term contracts are redundant but in an uncertain world with missing long-term futures markets they are essential for reducing finance costs.

penetration reaches a critical level. Storage or exports (for interconnected regions) can delay but not prevent the onset of the level of penetration beyond which curtailment (reducing VRE output) is needed.

Consider wind as the exemplar VRE whose output in any hour per MW of capacity is $\widetilde{\phi}$, a random variable independently drawn from its distribution for each hour, and thus with a constant expectation, ϕ. To simplify, its variable cost is taken as $v_W = 0$ (both assumptions are relaxed in Newbery (2020). If name-plate installed capacity of wind is W, its EFC is taken at δW, determined below. The model above is now modified by dropping nuclear power, and replacing the reliability constraint with $D(L) = P + F + \delta W$. Conventional plant now needs to deliver the residual demand – demand net of VRE, $R(t)$. The potential supply to meet total demand $D(t)$ is $G(t) + \phi(t)W$ (where $G(t)$ is output from conventional generation) but his may be excessive, in which case the wind will need to be curtailed.

In addition, there are a set of requirements to ensure system stability, explained in more detail in Newbery (2021a). For present purposes the relevant constraint is that the share of non-synchronous[11] generation (specifically VRE, as it is connected to the grid asynchronously and cannot normally deliver inertia) must be kept below a specified fraction of demand to reduce the rate at which frequency drops with a supply loss or a sudden increase in demand. The Grid Codes specify the allowable Rate of Change of Frequency (RoCoF) that determines the amount of inertia to avoid breaching the RoCoF standard.[12] This is normally specified by the maximum acceptable System Non-Synchronous Penetration (SNSP). Thus in the Single Electricity Market (SEM) of the island of Ireland studied in Newbery (2021a) the target 2020 SNSP is 75% (since achieved).

The level of SNSP will be critical in determining the amount of curtailment and hence the size of the resulting market distortion, and to that end define β = 1 - SNSP (so β = 25% in this case). Thus $G(t) \geq \beta(D(t)$ and curtailment will be needed in amount $k(\phi(t)W, t) = Max(0, \phi(t)W - (1 - \beta)D(t))$. While $\phi(h)W$ is potential wind output, actual or useful wind output will be $w(t) = \phi(t)W - k(\phi(t)W, t)$. Residual demand is then $R(t) = D(t) - w(t)$ and can be ordered for the set of hours with and without curtailment. Define h as hours without curtailment, ordered so that $R(h), R' < 0$, over $[0, H - h^*)$. For the remaining h^* hours wind is curtailed. It is convenient to define y as hours with curtailment with the curtailment function $k(\phi(y)W, y) \equiv k(y, W)$ separately ranked with $k' < 0$, over the range $[0, y^*)$, where $y^* = h^*$ is

[11] Non-synchronous generation is all plant without a spinning mass directly synchronized to the grid frequency (like wind, solar PV and DC interconnectors).

[12] Electrical equipment and synchronous generators automatically disconnect if they detect a higher than specified RoCoF for protection. If generation trips off as a result it would exacerbate the RoCoF and in a serious case might cause a black out, or at least require controlled disconnection, as happened in GB on 9 August 2019.

the solution to

$$k(y^*, W) = 0. \tag{4}$$

An example may be helpful. Suppose, quite plausibly,[13] that $D(h) = M - (M - m)h/H$ on $[0, H]$ and (implausibly) that $\phi(h)$ is linear and perfectly negatively correlated with demand: $\phi(h) = h/H$. Then $h^* = y^*$ solves

$$\begin{aligned} W(1 - h^*/H) &= (1 - \beta)\{M - (M - m)(1 - h^*/H)\}, \\ (1 - h^*/H) &= \frac{(1 - \beta)m}{W + (M - m)(1 - \beta)}. \end{aligned} \tag{5}$$

Figure 2 illustrates the residual demand curve, curtailment and actual wind output, plotted as functions of h as curtailment increases monotonically beyond $H - h^*$ because of its perfect negative correlation with demand.

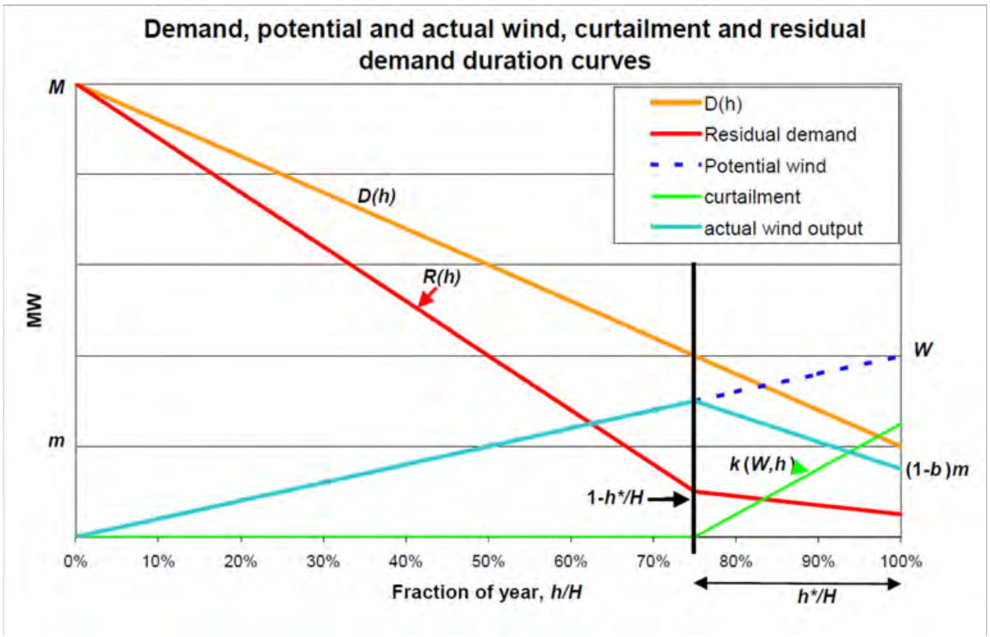

Figure 2: Hypothetical duration curves for perfectly negatively correlated wind

Figure 3 gives an illustrative (but still stylized) example using GB demand and actual wind data for 2018, but scaling wind up every hour by a factor of three, and then considering an increase in wind capacity of 1,000 MW (from the assumed start level of 39,100MW).[14] It shows the residual demand ranked in descending order over all hours, h, with the volume of wind

[13] The demand duration curve of Figure 1 was taken from GB data and is nearly linear over much of its length.

[14] GB demand is as measured, PV is ignored, actual wind in each hour is trebled, all storage and exports/imports are ignored.

curtailed in the same hour, h. The curtailment function is then graphed as $k(y, W)$ on $[0, y^*)$ with $k' < 0$. In this more realistic case where wind has little correlation with demand there is no simple relation between h and y.

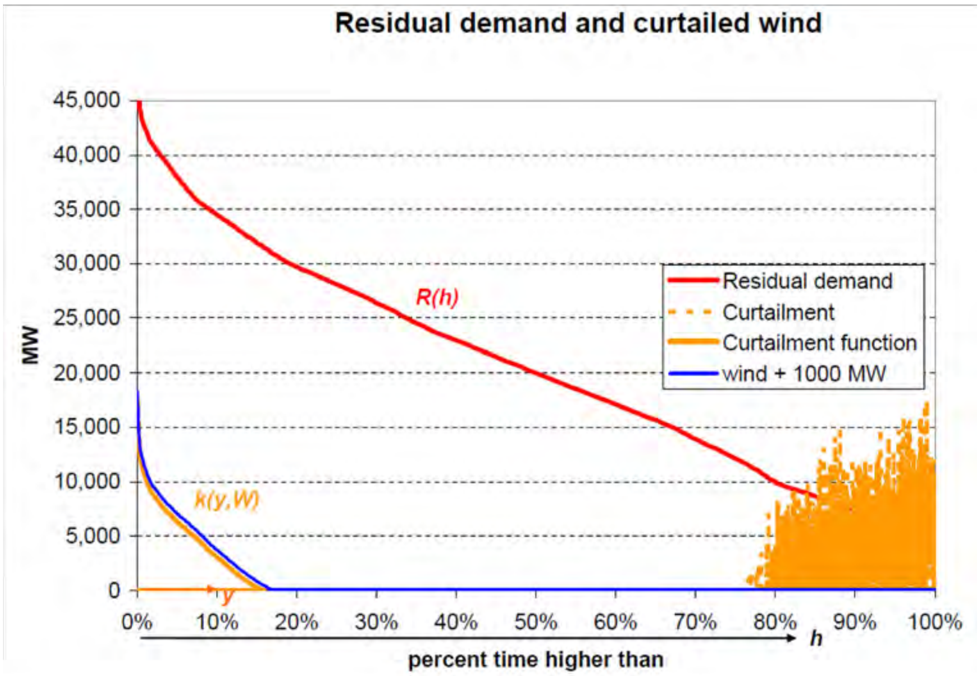

Figure 3: GB residual demand and curtailment function, scaled 2018 wind

The normal way to measure curtailment is the volume of wind curtailed, $\int_0^{h^*} k(W, h)dh$, which in general will be higher than $h^* W\phi$ as curtailment hours are likely to be hours of above average capacity factors. Existing wind farms experience average curtailment per MW of installed capacity (and the associated hours of zero profit) of $\int_0^{h^*} k(W, h)dh/W$ per MW of capacity. Marginal curtailment caused by the entry of 1 MW of extra wind capacity is

$$\begin{aligned} \frac{\partial}{\partial W}\int_0^{h^*} k(W,h)dh &= k(W,h^*)\frac{\partial h^*}{\partial W} + \int_0^{h^*} \frac{\partial k(W,h)}{\partial W}dh, \\ &= \int_0^{h^*} \frac{\partial k}{\partial W}dh. \end{aligned} \quad (6)$$

The ratio of the marginal to average curtailment is $W \int_0^{h*} \frac{\partial k}{\partial W}dh / \int_0^{h*} kdh$. In figure 3, the curtailment function is roughly linear in y over much of its range and can be approximated by

$$k = \alpha(W - W_0)(1 - y/y^*), \quad (7)$$

with W_0 the level of wind at which curtailment first appears. In this case $W_0 = 20,551$ MW, $W = 35,928$ MW),[15] $\alpha = 0.318$, and $y^* = h^* = 1,361$ hrs. Appendix A shows that the ratio of

[15] The peak capacity factor for the whole of the UK is 93.4% (as wind famrs in different locations are not perfectly

the marginal to average curtailment from (12) is just $2W/(W - W_0)$ or 4.7, considerably greater than 2. Newbery (2021a) gives more realistic estimates for island of Ireland in 2026 taking account of storage and exports and finds the ratio 3.66. For future reference, in the GB case above $\partial h^*/\partial W = 0.108$.

Total system costs with wind but without nuclear and replacing $P = D(L) - F - \delta W$, where δW is the EFC of wind capacity W, will be

$$\begin{aligned} C &= Wr_W + Fr_F + (D(L) - F - \delta W)(r_P + v_P L) + V\int_0^L (D(h) - D(L))dh \\ &\quad + v_P \int_L^{h_P} (D(h) - F - \phi W)dh + v_F F h_p + v_F \int_{h_P}^{H-h^*} (D(h) - \phi W)dh, \end{aligned} \tag{8}$$

as the variable cost of wind is zero. The first-order condition for minimizing fossil generation cost are unchanged:

$$\begin{aligned} 0 &= \frac{\partial C}{\partial F} = (r_F - r_P) - (v_P - v_F)h_P, \\ h_P &= \Delta r/\Delta v, \quad \Delta r \equiv r_F - r_P, \quad \Delta v \equiv v_P - v_F, \end{aligned}$$

and the length of time the peaking plant is needed is invariant to installed capacities (although they do depend on demand).

The total surplus (consumer surplus less generation cost) is $S = V\int_0^H (D(h)dh - C$, which, after noting that the envelope condition allows us to remove all terms in F from (8), becomes

$$\begin{aligned} S &= V\int_L^H D(h)dh - Wr_W - (D(L) - \delta W)(r_P + v_P L) \\ &\quad + v_P \int_L^{h_P} (\phi W - D(h))dh + v_F \int_{h_P}^{H-h^*} (\phi W - D(h))dh. \end{aligned} \tag{9}$$

5.1 Corrective tax on wind entry

Curtailment implies that the efficient price during curtailed periods will be the avoidable cost of wind (or VRE more generally), taken as zero, and as $\partial h^*/\partial W > 0$, additional wind will cannibalize the revenue from existing wind, as the number of profitable hours will decrease. However, new entrants enjoy the average, not the marginal curtailment that is relevant for assessing the benefits of additional wind investment. The benefit of an extra MW of wind capacity will be, from (9) and (3):

$$\begin{aligned} \frac{\partial S}{\partial W} &= \delta(r_P + v_P L) + \phi v_P (h_P - L) + \phi v_F (H - h^* - h_P) \\ &\quad - v_F (D(H - h^*) - \phi W)\partial h^*/\partial W - r_W, \\ &= (\delta V - \phi v_P)L + \phi\{r_F - r_P + v_F(H - h^*)\} - v_F \beta D(H - h^*)\partial h^*/\partial W - r_W. \end{aligned} \tag{10}$$

correlated) so the peak wind output is divided by 0.934 to derive the implied capacity. The increase in 1,000 MW gives an increase in peak wind output of 934 MW.

In the first line note that $h_P = \Delta r/\Delta v$ while in the second line $\phi W = (1-\beta)D(H-h^*)$, and so simplifies to $v_F \beta D(H-h^*)\partial h^*/\partial W$. The last line gives the surplus from 1 MW of extra wind, which can be compared to the market revenue below.

5.2 Decentralizing the efficient solution

As before, efficient prices, $p(h)$, are equal to the System Marginal Cost (if not load shedding) or the VoLL (when shedding load). For $0 \leq h \leq L$, $p(h) = V$, for $L < h \leq h_P$, $p(h) = v_P$, for $h_P < h < H - h^*$, $p(h) = v_F$, and for $H - h^* \leq h \leq H$, $p(h) = 0$ (the avoidable cost of wind). As before, free entry with consistent choices of V, L ensures conventional plant just covers its costs. The expected market unit net surplus (per MW of wind) (as $E\widetilde{\phi} = \phi$) will be

$$
\begin{aligned}
M_W &= \phi\{VL + v_P(h_P - L) + (H - h^* - h_P)v_F\} - r_W, \\
&= \phi\{r_P + h_P(v_P - v_F) + (H - h^*)v_F\} - r_W, \\
&= \phi\{r_F + (H - h^*)v_F\} - r_W.
\end{aligned}
$$

If it is left to wind producers to decide whether or not to enter,[16] then efficient entry requires that marginal surplus/MW, $\partial S/\partial W$, of equation (10) is equal to the expected net market surplus/MW. Normally one might expect that if all externalities (emissions pricing, learning spillovers) are internalized, then the efficient equilibrium ought to be supported in a competitive market, but that is not the case here. Instead it requires an annual fixed charge, τ/MW (if negative, a subsidy) to restore equality and hence efficient entry, with $\tau = M_W - \partial S/\partial W$:

$$
\begin{aligned}
\tau(W) &= \phi\{r_F + (H - h^*)v_F\} - r_W - (\delta V - \phi v_P)L - \phi\{r_F - r_P + v_F(H - h^*)\} \\
&\quad + v_F \beta D(H - h^*)\partial h^*/\partial W + r_W, \\
&= (\phi - \delta)VL + v_F \beta D(H - h^*)\partial h^*/\partial W > 0, \qquad (11)
\end{aligned}
$$

substituting for $\phi r_P = \phi(V - v_P)L$ in the top line. In the absence of any need to curtail wind, $h^* = \partial h^*/\partial W = 0$, and (11) can be interpreted as a method of de-rating wind to achieve efficient entry, $\delta = \phi$, consistent with the claim that in the absence of any correlation of wind with demand, wind should be de-rated by its average capacity factor, ϕ. Allowing for such correlations gives a different result (see Newbery, 2020). Otherwise, as $\partial h^*/\partial W > 0$, the tax needed for efficient entry is positive. The simplest way in which the systems charge could be levied is as an annual Transmission Network Use of Systems charge (TNUoS, to use the British terminology), which would depend on the expected level of curtailment measured by $h^*(W)$.

[16] Greve and Rocha (2020, p91) note that a 2019 Dutch off-shore wind tender "introduced a no subsidy requirement."

5.3 Numerical estimates

The corrective tax is best measured as a percent of the annual fixed cost, τ/r_W:

$$\tau(W)/r_W = (v_F/r_W)\beta D(H - h^*)\partial h^*/\partial W.$$

For the illustrative example of figure 3, $\beta = 25\%$, $\partial h^*/\partial W = 0.108$ and $\beta D(H - h^*) = 8,812$ hrs. Table 1, taken from Newbery (2020, Table 2) for the Single Electricity Market (SEM) of the island of Ireland, gives values for v_F (using 2019 prices) and r_F, giving $\tau/r_F = 48\%$, high because it ignored important relevant features of storage and export to avoid curtailment. More soundly based data from Newbery (2020) for the SEM considered two cases for β (25% and in the ambitious case, 15%) and for the export and storage opportunities in 2026. The costs are converted (at 2018 exchange rates of €1.13=£1) to € and shown in Table 1. The projected median gas price is €21.4/MWh (FES, 2019) while the CO_2 price is taken as €40/tonne. The corrective charge in the first case ($\beta = 25\%$) is $\tau/r_W = 20\%$ and in the ambitious case just under 10% of annual fixed costs.

Table 1 Cost estimates[17]

r_F	€85,218/MWyr	v_F	€61/MWh
r_P	€37,012/MWyr	v_P	€91/MWh
r_W	€120,132/MWyr	v_W	€7/MWh
Δr	€48,206/MWyr	Δv	€30/MWh
h_P	1,607 hours	L	8 hours

5.4 Learning externalities

The assumption above was that the learning spillovers were already corrected, but the empirical figures for the annualized capital costs were not so corrected. Newbery (2018) shows how to calculate the globally desirable level of subsidy and Newbery (2020) derives the values, with a central estimate for the SEM in 2026 of 10% of the capital cost. This is comparable to, and offsets, the ambitious scenario corrective charge and therefore roughly cancels it out. Taking the uncorrected IRENA (2016) learning rate estimates at face value, the learning subsidy might be 16% of the capital cost, again, not far short of the corrective charge in the base case, at least for the SEM.

[17] BEIS (2020) gives 2025 (medium) capital costs for base and peaking plant (open-cycle gas turbine) and on-shore wind, as well as the fixed and variable operating costs and fuel efficiencies. The capital cost figures for base and peaking are per *derated* MW, and so the cost per installed MW needs to be inflated to allow for this.

6 Conclusions

Once a wind turbine is commissioned and connected, it will generate so long as it is not constrained or off-line. Some constraints are local, caused by transmission limits,[18] and are best handled by nodal pricing (as in the U.S. Standard Market Design) or offering non-firm connections in such locations. The constraints considered here are system wide, and need a system-wide solution. The first part of good system design is to ensure that carbon costs are properly charged, innovative technologies are compensated for their external learning benefits, and electricity pricing into the grid reflects the system marginal cost of generation, cleansed of distortionary subsidies.

The remaining element of good design is to ensure the efficient entry (and type) of new generation. With an efficient (perfect foresight) energy-only market, or suitably auctioned capacity payments (shown to be equivalent in Newbery, 2020), fossil entry can be left to market signals. The capacity credit for wind is rather more complicated to calculate (and very sensitive to demand and wind conditions in winter months, as Newbery (2020) shows). The key new factor considered here is that once wind penetration is high enough to cause system-wide curtailment, additional wind imposes an additional cost that is not reflected in market prices, as the marginal curtailment is many times higher than the average curtailment that sets prices. This is the "tragedy of the commons" that is at the heart of the market failure. These extra costs are almost proportional to $\beta = 1-$ SNSP, but will also be affected by the amount of storage and the ability to export, and seem to be material. Export opportunities in turn depend on export capacity but also on the extent and simultaneity of wind abroad. Offsetting this corrective charge, the global learning externality (mostly reaped abroad, but internalized if other countries offer similar subsidies as a club payment, e.g. under the EU *Clean Energy Package*) might be 11-17% of annual fixed costs and therefore of comparable magnitude under favourable circumstances.

The conclusion is that the capacity credit might need separate attention and that the curtailment effect will depend very much on system characteristics (penetration and SNSP most directly) and is comparable to or larger than the likely justified global learning subsidy. Whether this would be true in other systems or with higher wind penetration should be explored as part of wider study of the appropriate way to procure wind, and the extent to and manner in which to grant capacity payments to wind. The simpler alternative is to set the renewables target and run auctions for the amount of renewables by allowing them to bid for the strike price in a Contract for Difference for the first 25,000 full operating hours (i.e. MWh/MW), which would provide a

[18] Massive VRE entry is likely to exacerbate congestion, unless transmission investement is coordinated and/or VRE is directed to places with adequate transmission capacity. See LaRiviere and Lyu (2022) for an interesting case study in Texas.

revenue stream for about 10 years. Recent Continental auctions for off-shore wind suggest that the strike price might be at or below that of conventional plant (Greve and Rocha, 2020).

References

Buchanan, J.M., 1965. An Economic Theory of Clubs, *Economica*, 32(125), 1-14.

BEIS, 2020. *ELECTRICITY GENERATION COSTS,* athttps://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/911817/electricity-generation-cost-report-2020.pdf

Csereklyei, Z., S. Qu, T. Ancev, 2019. The effect of wind and solar power generation on wholesale electricity prices in Australia, *Energy Policy*, 131, 358-369, https://doi.org/10.1016/j.enpol.2019.04.007.

Clò, S., A. Cataldi, P. Zoppoli, 2015. The merit-order effect in the Italian power market: The impact of solar and wind generation on national wholesale electricity prices, *Energy Policy*, 77, 79-88, https://doi.org/10.1016/j.enpol.2014.11.038.

Cludius, J., H. Hermann, F. Chr. Matthes, V. Graichen, 2014. The merit order effect of wind and photovoltaic electricity generation in Germany 2008–2016: Estimation and distributional implications, *Energy Economics*, 44, 302-313, https://doi.org/10.1016/j.eneco.2014.04.020.

Deane, P. S. Collins, B. Ó. Gallachóir, C. Eid, R. Hartel, D. Keles and W. Fichtner, 2017. Impact on Electricity Markets: Merit Order Effect of Renewable Energies, Ch 16 in *Europe's Energy Transition - Insights for Policy Making*, Academic Press, 2017, https://doi.org/10.1016/B978-0-12-809806-6.00016-X , also at http://www.innoenergy.com/wp-content/uploads/2016/03/RREB_III-Merit_order_Final.pdf

FES, 2019. *National Grid ESO Future Energy Scenarios 2019 Data Workbook*, at https://www.nationalgrideso.com/future-energy/future-energy-scenarios/fes-2019-documents.

Green, R.J. and T-O. Léautier, (2015). "Do costs fall faster than revenues? Dynamics of renewables entry into electricity markets", Toulouse School of Economics TSE-591 at http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.697.4750&rep=rep1&type=pdf

Green, R. & Vasilakos, N., 2010. Market behaviour with large amounts of intermittent generation. *Energy Policy* 38, 3211–3220.

Greve, T. and M. Rocha, 2020. Policy and theoretical implications of zero-subsidy bids in the German offshore wind tenders, *The Energy Journal*, 41(4), 89-104.

Hardin, G., 1968. The Tragedy of the Commons, *Science*, 162, Issue 3859, 1243-1248. DOI: 10.1126/science.162.3859.1243

Heptonstall, P.J. and R. J. K. Gross, 2020. A systematic review of the costs and impacts of integrating variable renewables into power grids. *Nature Energy*, at https://doi.org/10.1038/s41560-020-00695-4.

Hill, A., 2021. Excessive entry and investment in deregulated markets: Evidence from the electricity sector, *Journal of Environmental Economics and Management,* 110 102543 p 1-21.

DOI: https://doi.org/10.1016/j.jeem.2021.102543

IRENA, 2019. *Future of wind: Deployment, investment, technology, grid integration and socio-economic aspects (A Global Energy Transformation paper)*, International Renewable Energy Agency, Abu Dhabi. At https://www.irena.org/wind

Joskow, P.L. and J. Tirole, 2007. Reliability and Competitive Electricity Markets, *RAND Journal of Economics*, 38(1): 68-84. At http://dx.doi.org/10.1111/j.1756-2171.2007.tb00044.x

Keane, A., M. Milligan, C. J. Dent, B. Hasche, C. D'Annunzio, K. Dragoon, H. Holttinen, N. Samaan, L. Soder and M. O'Malley, 2011. Capacity Value of Wind Power. *IEEE Transactions on Power Systems* 26(2): 564–572. https://doi.org/10.1109/TPWRS.2010.2062543

Ketterer, J.C., 2014. The impact of wind power generation on the electricity price in Germany, *Energy Economics*, 44, 270—280.

LaRiviere, J. and X. Lyu. 2022. Transmission constraints, intermittent renewables and welfare. *Journal of Environmental Economics and Management,* 112, 102618, p1-23. DOI: https://doi.org/10.1016/j.jeem.2022.102618

National Grid, 2014. *EMR Electricity Capacity Report*, June.

Newbery, D., 2018. Evaluating the case for supporting renewable electricity, *Energy Policy*, 120, 684–696. https://doi.org/10.1016/j.enpol.2018.05.029

Newbery, D. 2020. Club goods and a tragedy of the commons: the Clean Energy Package and wind curtailment, EPRG WP 2036 at https://www.eprg.group.cam.ac.uk/eprg-working-paper-2036/

Newbery, D., 2021a. Implications of the *National Energy and Climate Plans* for the Single Electricity Market of the island of Ireland, *Energy Policy* 158, 112513, 1-11, at

https://doi.org/10.1016/j.enpol.2021.112513.

Newbery, D., 2021b. Designing efficient Renewable Electricity Support Schemes, EPRG WP 2107 at https://www.eprg.group.cam.ac.uk/eprg-working-paper-2107

Pudjianto, D., Aunedi, M., Djapic, P., Strbac G., 2014. Whole-Systems Assessment of the Value of Energy Storage in Low-Carbon Electricity Systems, I*EEE Transaction on Smart Grid*, 5, 1098-1109.

Stahn, H. and A. Tomin, 2021. Externality and common-pool resources: The case of artesian aquifers. *Journal of Environmental Economics and Management* 109, 102493, p1-17 DOI: https://doi.org/10.1016/j.jeem.2021.102493

Starrett, D., 1972. Fundamental nonconvexities in the theory of externalities, *Journal of Economic Theory*, 4(2), 180-199. https://doi.org/10.1016/0022-0531(72)90148-2

Stoft, S., 2002. *Power System Economics: Designing Markets for Electricity*, Wiley.

Ueckerdt, F., L. Hirth, G. Luderer, O. Edenhofer, 2013. System LCOE: What are the costs of variable renewables? *Energy*, 2013, vol. 63, issue C, 61-75. At https://www.pik-potsdam.de/members/edenh/publications-1/SystemLCOE.pdf

Weiss, T. and A. Wänn, 2013. Does Ireland need more storage? Power system overview and RES integration, at

https://www.store-project.eu/documents/target-country-results/en_GB/energy-storage-needs-in-ireland-executive-summary

Zachary, S., A. Wilson and C. Dent, 2019. The integration of variable generation and storage into electricity capacity markets, at https://arxiv.org/abs/1907.05973

Appendix A Linearizing curtailment functions

The curtailment function in figure 3 is roughly linear and can be approximated by (now replacing y by h for convenience)

$$k = A(1 - h/h_r^*) + \alpha(W - W_r),$$

where subscript r refers to a reference level of wind, W_r. If this holds over a wide enough range then there will be no curtailment until $W = W_0$, in which case $A = \alpha(W_r - W_0)$, and

$$\begin{aligned}
h^* &= h_r^* \frac{W - W_0}{W_r - W_0}, \\
k &= \alpha(W - W_0)(1 - h/h^*), \\
\frac{\partial h^*}{\partial W} &= \frac{h_r^*}{W_r - W_0} = \frac{h^*}{W - W_0}.
\end{aligned}$$

It follows that

$$\begin{aligned}
\int_0^{h*} k dh &= \frac{1}{2}\alpha(W - W_0)h^*, \\
\int_0^{h*} \frac{\partial k}{\partial W} dh &= \alpha \int_0^{h*} (1 - h/h^*) dh + \alpha(W - W_0) \int_0^{h*} \frac{h}{h^{*2}} \frac{\partial h^*}{\partial W}, \\
&= \alpha h^*.
\end{aligned}$$

The ratio of the marginal to the average curtailment is then

$$\frac{W \int_0^{h*} \frac{\partial k}{\partial W} dh}{\int_0^{h*} k dh} = \frac{2W}{W - W_0} > 2. \tag{12}$$