Long-run Equilibrium Effects of Variable Renewable Energy*

P. Holmberg†   A. B. Philpott‡

April 3, 2026

Abstract

This paper studies the long-run equilibrium effects of variable renewable electricity (VRE) in a competitive power market with free entry and a continuum of thermal generation technologies. The analysis shows that the entire equilibrium price distribution and the loss of load probability both remain unchanged under moderate VRE entry. However, thermal investment shifts away from fuel-efficient baseload technologies toward less fuel-efficient technologies. We identify circumstances where moderate VRE entry increases total fuel costs and emissions. The finding is particularly relevant for power systems with low demand variability, negligible hydropower capacity, and no domestic cap on carbon emissions. Sufficiently high VRE penetration reduces fuel costs and emissions, and makes the price distribution more favourable for energy-intensive industries.

Keywords: Investment, electricity market, variable renewable energy, long-run equilibrium, operating costs, technology mix, CO_2 emissions

JEL Classifications: D41, L94, Q41, Q42, Q48, Q54

*We are grateful for comments from Lina Reichenberg. The authors thank the Institute for Mathematical and Statistical Innovation (IMSI) at the University of Chicago for support and hospitality during the long program *The Architecture of Green Energy Systems,* where parts of this work were done. Pär Holmberg has been financially supported by the Swedish Energy Agency (project P2022-00667). We thank Lisa Brogårdh for research assistance.

†Corresponding author: Research Institute of Industrial Economics (IFN), Box 55665, Stockholm, Sweden. E-mail: par.holmberg@ifn.se. Phone: +46 8 6654500. Associate Researcher of the Energy Policy Research Group (EPRG), University of Cambridge, and the Program on Energy and Sustainable Development (PESD) at Stanford University.

‡Faculty of Engineering and Design, University of Auckland, New Zealand. E-mail: a.philpott@auckland.ac.nz.

1 Introduction

At the 28th Conference of the Parties of the UNFCCC (COP28) in 2023,[1] 130 countries pledged to triple the world's renewable-energy capacity by 2030 (REN21, 2024). According to Borenstein (2012), unpriced pollution externalities are the primary public policy argument for subsidising renewable electricity generation, including Variable Renewable Energy (VRE) such as solar and wind power.[2] Borenstein (2012) finds that other common arguments, such as energy security, are unlikely to substantially alter the analysis. But the EU, in directives 2009/28/EC, 2018/2001 (RED II), and 2023/2413 (RED III), partly justifies support for renewable energy on the grounds of securing energy independence and improved industrial competitiveness.[3]

Short-term studies of VRE entry have confirmed some of these policy rationales, particularly the effects on electricity prices and emissions. Average electricity prices tend to decrease due to the *merit-order effect* (Sensfußet al., 2008; Hirth, 2013; Liski and Vehviläinen, 2020; Reichenberg et al., 2023). Moreover, Graf and Marcantonini (2017) find that a 10% increase in photovoltaics and wind infeed in Italy reduces the average thermal plant's yearly CO_2 emissions by about 2%.

We characterise a long-run competitive equilibrium (Antweiler and Muesgens, 2021; Holmberg and Ritz, 2021; Newbery, 2016). This is an equilibrium where the capacities of all technologies have had time to adjust to the entry of VRE. We abstract from emission caps, hydropower, and demand response, though the marginal cost of thermal production may include a fixed carbon price.

We identify circumstances where moderate VRE entry is not effective in reducing fuel dependency. On the contrary, total fuel costs and CO_2 emissions may increase as VRE capacity increases, especially in countries with baseload technologies that have near-zero marginal costs, such as nuclear power and geothermal energy. The reason is that a large share of baseload capacity in the generation mix is no longer cost effective when VRE capacity increases (Antweiler and Muesgens, 2021; Lamont, 2008; Newbery, 2016). Instead, it becomes cost effective to invest in technologies with low capital cost but poor

[1] The United Nations Framework Convention on Climate Change (UNFCCC) is the UN framework for negotiating global climate agreements.

[2] Wave energy, tidal energy, ocean current energy, and run-of-river hydropower are other examples of VRE.

[3] Moreover, the EU argues that renewable-energy investments stimulate technological development with positive spillovers. Borenstein (2012) considers spillover arguments persuasive for subsidising basic energy research, but much less so for subsidising the deployment of renewable technologies. Our study does not evaluate this policy aspect.

fuel efficiency.[4]

The shift towards less fuel-efficient technologies is particularly strong when the variability of demand is small compared to the variability of the VRE output. In this case, before VRE entry, the thermal generation mix is dominated by fuel-efficient baseload (nuclear or geothermal energy). High penetration of VRE is needed, beyond the threshold at which all baseload technologies with zero marginal costs exit the market, before fuel costs and CO_2 emissions eventually start to decline. We argue that the risk of increased CO_2 emissions and increased fuel dependency due to VRE entry is particularly relevant in jurisdictions with no hydropower and no domestic carbon cap.

We find that there is no long-run reduction in average electricity prices under moderate VRE entry, in line with Antweiler and Muesgens (2021).[5] Hence, energy-intensive industries, which tend to have constant consumption throughout the year, do not benefit from moderate VRE entry.[6] Other consumers whose demand varies during the year will normally spend less on electricity. The reason is that the correlation between demand and price decreases after VRE entry, as long as it is moderate.

VRE entry can increase price volatility in the short run (Ketterer, 2014). But according to our long-run model, volatility does not change as long as entry is moderate. High penetration of VRE increases volatility. This does not mean that price spikes become higher or more frequent. Instead, zero-price outcomes become more frequent. Energy-intensive industries prefer the price distribution under high VRE penetration to the distribution before VRE entry, even if these industries were highly risk-averse. High penetration of VRE leads to lower average electricity prices, which is consistent with Newbery (2016). Thus, there is a long-run merit-order effect in this case.

[4]In addition, more variable output decreases the efficiency of a given plant, due to ramping costs (Gutiérrez-Martín et al., 2013; Valentino et al., 2012). This *load-cycling effect* increases carbon emissions per energy produced. Graf and Marcantonini (2017) show that a 10% increase in photovoltaics and wind infeed in Italy increases the average plant's emissions relative to its output by about 0.3%. Kaffine et al. (2020) find that the load-cycling effect diminishes the reduction of CO_2 emissions in the Southwest Power Pool in the U.S. by 6.5%. According to the analysis by Cullen and Reynolds (2023), accounting for the load-cycling effect could increase the cost of CO_2 reduction by 40-80% in Texas. The effect of load cycling should be smaller in today's and future power systems, where batteries and load shifting are relatively cheap. We have neglected the load-cycling effect. Our model assumes flexible thermal plants without ramping costs.

[5]In an extension, Antweiler and Muesgens (2021) show that, for the case where baseload is provided by a monopolist, the average price would go down, also in the long run, when VRE enters the market.

[6]Rationed energy may change with entry, but in the model consumers are indifferent at the margin because served demand pays the reservation price when rationing occurs.

Our method is related to peak-load pricing (Boiteux, 1960; Turvey, 1968; Crew and Kleindorfer, 1986; Stoft, 2002; Biggar and Hesamzadeh, 2014; Léautier, 2019).[7] This method is used to determine optimal investments in a system without storage so that goods have to be delivered just in time, which is approximately the case in many electric power systems (Chao, 1983, 2011).[8] The optimal mix typically includes efficient baseload, which has a high investment cost and is suitable for continuous operation, and inefficient peakers, which have a low investment cost and are only operated when there is a shortage of power. Mid-merit plants, which fall in between these two extremes, typically run when demand is moderate or high. In practice, baseload, mid-merit and peakers are often thermal power technologies that use some sort of fuel – uranium, coal, gas or oil – to produce electricity.

Our model of thermal plants is similar to Holmberg and Ritz (2021). We allow investors to choose from a continuum of thermal generation technologies in a market with demand that is inelastic up to some reservation price.[9] The reservation price is referred to as the *Value of Lost Load* (VOLL), which is assumed to equal consumers' marginal value of avoided disconnection. The continuum of thermal plants has an operating (marginal) cost per unit of energy from zero upward.[10] Each thermal technology is indexed by its marginal cost c. We abstract from a carbon cap. The function $k(c)$ gives the investment cost per unit of capacity of technology c. We assume that fuel prices are constant and independent of investments. Moreover, we assume that technologies with zero marginal cost have zero emissions and that technologies with a positive marginal cost have positive CO_2 emissions. Hence, the analysis does not cover systems with large-scale use of biofuels or electrofuels in thermal plants, such as Sweden, where thermal generation is essentially

[7] Willems and Yu (2023) study how efficiency and investments would change if the spot market implemented discriminatory pricing. Zöttl (2010) uses a peak-load pricing model to study how investments are influenced by market power. Eisenack and Mier (2019) extend the peak-load pricing model by considering that different production technologies can be adjusted within their capacity at different speeds.

[8] Chao (1983, 2011) outlines efficient pricing in markets with supply uncertainty and intermittent resources.

[9] The reservation price could be set arbitrarily high if one wants to neglect this technicality. Increasing the reservation price influences investment in peakers with high marginal costs, but the capacities of other technologies do not depend on the reservation price (Holmberg and Ritz, 2021).

Similarly, considering batteries and load shifting should mainly influence investments in peakers. We believe that demand elasticity at low prices would have a larger impact on investments in baseload and mid-merit. This is suggested, for example, by results in Green and Vasilakos (2011) and Borenstein (2005).

[10] In practice, the fuel cost of thermal plants is not zero, but the fuel cost of nuclear power and geothermal energy is very low and close to zero.

fossil free.

Each unit of VRE capacity may consist of solar power, wind power or a mix of VRE technologies. It can be helpful to think of each VRE unit as representing renewable generation capacity that is spread across the region under consideration, so that any two VRE units are essentially indistinguishable. Thus, we assume that VRE units can be replicated, allowing VRE capacity to expand at constant returns to scale. We allow demand and VRE output to be correlated and assume that the marginal cost of VRE is zero.

We solve for a long-run equilibrium, where thermal power investments are optimized given VRE entry, or where investments in thermal power and VRE are optimized jointly. Entry of VRE could be driven by technological development, policy targets or subsidies, but (similar to much of the literature on VRE entry) we abstract from the financing side of support schemes and how they influence market participants.[11] We examine how the operation and investment decisions for thermal power plants, and their emissions, are affected by the entry of VRE.

In the long-run equilibrium, it is cost effective to use a thermal technology with marginal cost c for plants that run with a specific probability, denoted $\pi(c)$. If this probability were higher, it would be optimal to switch to a more fuel-efficient technology with a lower operating cost. Conversely, if the probability were lower than $\pi(c)$, it would be more cost effective to switch to a less fuel-efficient technology with a lower investment cost. Hence, $\pi(c)$ is decreasing with respect to c. In a cost-optimal configuration, the probabilities $\pi(c)$ depend only on the trade-off between investment and operating costs, which is consistent with the screening-curve analysis (Newbery, 2016). The probabilities $\pi(c)$ do not depend on anything else, including the probability distribution of demand net of VRE output, provided that it remains optimal for the technology with cost c to partly stay in the market.

We consider a perfectly competitive market. Thus, the probability $\pi(c)$ also reflects the probability that the price exceeds c.[12] As long as technology c remains active in the market, the distribution of prices above c is independent of the distribution of net demand. We define VRE entry as moderate when the probability that VRE output exceeds demand is less than $1-\pi(0)$, so that some baseload with $c = 0$ runs with a probability higher than $\pi(0)$. Baseload capacity with zero marginal cost will gradually exit the market when VRE

[11] An exception is Newbery (2016) who analyses how the fossil cost savings from VRE entry should be allocated between electricity consumers and general taxation.

Brown and Reichenberg (2021) study how different policy instruments for promoting VRE affect investment incentives in a long-run equilibrium.

[12] In a competitive market, a plant runs if and only if the price is above its marginal cost.

enters, but no thermal technology will completely exit the market as long as VRE entry is moderate. There might be outcomes where VRE output is above demand so that VRE is curtailed, but overall the probability of having zero prices does not increase. Under these moderate conditions, the distribution of prices (and the average price) remains unchanged as VRE penetration increases.

For low penetration of VRE, low prices occur when demand is low, but with medium penetration of VRE, low prices increasingly occur when the VRE output is high. This transition strengthens the negative correlation between prices and VRE output. Hence, VRE producers receive a lower average price than the market average. This effect is often referred to as the *cannibalization effect* (Hirth, 2013; Prol et al., 2020; Reichenberg et al., 2023). We show that the effect strengthens as VRE capacity increases and continues to intensify when entry is high and VRE is frequently curtailed, which is consistent with Hirth (2013) and Reichenberg et al. (2023).[13] Entry of VRE implies that the market has to invest in more flexibility, which corresponds to thermal power with high marginal cost in our model. Additional flexibility is costly (Hirth et al., 2015). The cannibalization effect ensures that VRE technologies pay for this via a lower output-weighted average price.

Many consumers buy more electricity during the day than during the night. Moreover, consumers far from the equator tend to buy more electricity during the winter, for heating, whereas consumers in the subtropics tend to buy more electricity during the summer, to cool the house. Thus, especially before the entry of VRE, the price tends to be high when consumers have high demand, and vice versa. Due to the positive correlation between demand and price, the demand-weighted average price is higher than the (unweighted) average price. This profile effect ensures that consumers pay for some of the flexibility in the grid. We show that this effect weakens as VRE enters the market, because high prices increasingly occur when VRE output is low. Hence, the demand-weighted average price decreases as VRE enters the market, even if the unweighted price is constant (for moderate entry).[14] Many energy-intensive industries essentially have constant demand

[13] The cannibalization effect may not be a problem for VRE investments in a power system with a carbon cap. Brown and Reichenberg (2021) demonstrate that capping emissions leads to higher electricity prices, which compensates the effect.

[14] Green and Léautier (2015) find related results in their simulation of a future British power system. The average price remains at £81/MWh when more wind power enters the system, but the demand-weighted average price is reduced from £87/MWh to £85/MWh. In their case, this is partly driven by a positive correlation between wind power and demand. Our model shows that there is such a weakening of the profile effect after VRE entry, even if VRE output and demand would be negatively correlated.

(flat demand) all through the year, so they pay the unweighted price.

We say that VRE capacity x is high when $G(x, 0)$, the probability that VRE output exceeds demand, is larger than $1-\pi(0)$. In this case, all thermal technologies with marginal cost c such that $G(x, 0) \geq 1 - \pi(c)$ exit the market entirely. We denote the highest such marginal cost by $\underline{c}\,(x)$. The total probability that the price lies in some fixed range $[0, c]$ remains unchanged after VRE entry, but under high VRE penetration where $\underline{c}\,(x) > 0$, all probability mass in the range $[0, \underline{c}\,(x)]$ collapses to a point mass at zero. This strengthens the cannibalization effect of VRE and reduces the average price. Thus there is a long-run merit-order effect for high penetration of VRE. In the long-run equilibrium, the distribution of prices above $\underline{c}$ is not influenced by entry of VRE, although the states of the world associated with high prices may change. Thus, prices for low levels of VRE entry first-order stochastically dominate prices for high VRE levels, which have more probability mass at the price zero. Irrespective of their risk preferences, consumers with flat demand, such as energy-intensive industries, prefer prices in a system with high levels of VRE entry.

For the extreme case where consumers' total demand is a constant, moderate entry of VRE actually leads to increased total operating costs (of thermal power) and increased CO_2 emissions. The explanation is that fuel-efficient baseload production with low c is replaced by VRE and fuel-inefficient thermal production with high c. For constant total demand, the substitution effect is so strong that fuel consumption and emissions increase even if the total output from thermal production decreases.

The shift to less fuel-efficient technologies is particularly damaging when it occurs for mid-merit technologies, which have both a relatively large dispatch probability and operating cost. We show that this is the case when the investment cost $k\,(c)$ is highly convex for technologies in the mid range and when intermediate renewable output levels are relatively rare. With constant demand, VRE entry will not begin to reduce total fuel costs until beyond the threshold where baseload has completely exited the market.

Entry of VRE increases investments in thermal power with high marginal costs (peakers), but it is socially inefficient to invest in technologies with an even higher marginal cost, above some cutoff $\overline{c}$. Above this cutoff it is more efficient to ration consumers, and have a partial blackout, instead of adding more production capacity (Chao, 1983; 2011; Joskow and Tirole, 2007). Similar to Holmberg and Ritz (2021), we show that, in the long run, this cutoff and the risk of a partial blackout, the Loss of Load Probability (LOLP), are constant irrespective of the probability distribution of demand and the level of VRE entry. In addition, we show that this is also the case when demand and VRE output are correlated and also for high levels of VRE entry (up to the point where VRE entry is so high that all thermal power technologies have completely exited the market). Even if LOLP is unchanged, the Expected Energy Not Served (EENS) can increase with more VRE in the system, if rationed energy increases once a partial blackout occurs. We identify circumstances when this is the case. This contribution can be of relevance for the development of reliability standards and the design of capacity mechanisms.

Our approach is analytical, which allows us to draw conclusions across a wide range of input parameters. Many of our conclusions are qualitative. An exception is the distribution of prices, for which we obtain a sharp prediction under fixed input parameters. The analytical perspective is particularly useful when discussing energy policies that are intended to persist over a long time or to apply across multiple jurisdictions. Our approach is closest to Antweiler and Muesgens (2021) and Newbery (2016), but both focus on a setting with two thermal technologies and derive results for specific distributional assumptions on demand and VRE output. Moreover, they consider fossil-based baseload in which case VRE entry is more likely to reduce fuel costs (Newbery, 2016).

Green and Vasilakos (2011) simulate the electricity market in the UK and find that, in the long run, wind-power investments lead to more thermal peakers and that price distributions are fairly stable with respect to wind entry. Green and Léautier (2015) show in a model with linear demand and a finite set of technologies that the expected prices on vertical segments of the supply curve remain unchanged by renewable entry. Their simulations (for high VRE levels) of the UK show nuclear exit and higher CO_2 emissions unless even more VRE is added. Empirically, Bushnell and Novan (2021) find that in California, VRE entry reduced the profits of baseload generators but not of peakers – consistent with our model's predictions.

The paper contributes to the study of the long-run equilibrium effects of additional VRE, especially in a power system with no domestic emissions cap, negligible hydropower and where baseload has approximately zero marginal cost. Examples of such jurisdictions are Belarus, Belgium, Czechia, Hungary, and the United Arab Emirates as well as the U.S. states Georgia, Nevada, North Carolina, and South Carolina.[15] The paper's main contributions to the study of long-run effects in such a power system can be summarised as follows:

1. Extending previous analytical results, derived for two thermal technolo-

[15] Many EU countries have committed to domestic emissions targets. VRE entry in an individual EU country may increase domestic emissions, even if there is an EU-wide cap. All else equal, this risk is higher in small EU countries, where domestic emissions have negligible influence on the EU-wide carbon price.

gies under specific distributional assumptions (Antweiler and Muesgens, 2021; Newbery, 2016) for a power system where baseload is fossil-based, to a continuum of thermal technologies, including baseload technologies with zero marginal cost, and an arbitrary joint distribution of demand and VRE output.

2. Detailed and general characterisation of the price distribution and its invariance, including when there is high entry of VRE.

3. Providing an analytical and general characterisation of circumstances where moderate VRE entry increases total fuel costs and CO_2 emissions.

4. Irrespective of correlation between demand and VRE output, we find that the positive correlation between demand and price generally weakens for moderate VRE entry, so that consumers with variable demand pay less on average even if the unweighted average price is unchanged.

5. Showing that risk-averse energy-intensive industries with flat demand generally benefit from high entry of VRE.

6. Showing that the cannibalization effect is in general monotonically strengthened when VRE enters the system, irrespective of correlation between demand and VRE output.

7. Generalising previous results on the invariance of the loss of load probability (Holmberg and Ritz, 2021) with regard to high entry of VRE and correlation between VRE output and demand. We also identify when Expected Energy Not Served (EENS) increases after VRE entry.

The remainder of the paper is organised as follows. Our model is introduced in Section 2. The long-run competitive equilibrium of the electricity market is characterised in Section 3. Section 4 studies the special case with constant demand. Section 5 concludes the paper. All proofs and some supporting technical lemmas are collected in the Appendix.

2 Model

Similar to Holmberg and Ritz (2021), we consider a two-stage competitive equilibrium model with sequential rationality. Investments occur in Stage 1 and production in Stage 2. The investment stage has free entry and the spot market in Stage 2 is perfectly competitive. There is a range of thermal technologies with a continuum of marginal costs. We let $k(c) > 0$ be the investment cost per unit of capacity for a technology with marginal cost $c \geq 0$. As in Holmberg and Ritz (2021), we assume that $k(c)$ is strictly convex, twice differentiable and satisfies $k'(c) \in (-1, 0)$ for positive c.[16] These properties make sure that each technology is on the efficient frontier of thermal technologies.[17] For theoretical completeness, we assume $\lim_{c\to\infty} k'(c) = 0$, though only the shape of $k(c)$ for $c < V$ affects our results.[18] At $c = 0$, we interpret $k'(0)$ as the right-hand derivative, since $c \geq 0$.

However, it is not self-evident that any thermal investments occur if the capacity of VRE is very high. We let $\underline{c} \geq 0$ be the lowest marginal cost and $\overline{c} \geq 0$ be the highest marginal cost for which thermal investments are cost effective. These cutoffs may depend on the capacity of VRE.

V is the reservation price (Value of Lost Load, VOLL) for consumers. Similar to Holmberg and Ritz (2021), we assume that thermal investment costs are sufficiently low and that consumers have a sufficiently high willingness to pay for electricity, so that

$$k(0) < -k'(0)\,V, \tag{1}$$

which will ensure that $\overline{c} > 0$.

Thermal technologies are assumed to be flexible, and without ramping costs. The cumulative installed capacity of thermal technologies with a marginal cost at or below c is denoted by $q(c)$, and provides the competitive supply curve in Stage 2. Negative investments are not allowed for any technology c, so $q(c)$ is non-decreasing. We assume that $q(c)$ is continuous from the right, which matters at $c = 0$. Note that the supply function describes the chosen thermal technology mix.

We allow demand and VRE output to be correlated. We define a bivariate random variable $(d, w) \in [0, \overline{d}] \times [0, 1]$, with joint bounded and continuous density $f(d, w)$ and distribution $F(d, w)$, where d denotes a realisation of inelastic demand and w a realisation of renewable availability. We assume

[16] Related assumptions are also made by Zöttl (2010) and Léautier (2019).

[17] The total expected cost of a unit that is dispatched with probability π is given by $\Pi = k(c) + \pi c$. If $k'(c) \geq 0$ for $c \in (c_1, c_2)$, then a technology with $c = c_1$ would have a strictly lower total cost than any technology in the interval (c_1, c_2). If $k'(c) < -1$ for $c \in (c_1, c_2)$, then a technology with $c = c_2$ would have a strictly lower total cost than any technology in the interval (c_1, c_2). Our assumptions rule out such cases. Strict convexity ensures that each thermal technology is uniquely optimal for some dispatch probability.

[18] As long as $k(c) \geq 0$, the shape of $k(c)$ for $c \geq V$ does not matter for our results, as it will be more cost efficient to ration demand. Still, the assumption $\lim_{c\to\infty} k'(c) = 0$ simplifies some of our arguments. It ensures that, conditional on a thermal technology being used, there is a uniquely optimal thermal technology also for low dispatch probabilities (capacity factors).

that $f(d,w) > 0$ in the interior of $[0,\overline{d}] \times [0,1]$ and set $f(d,w) = 0$ outside the region $[0,\overline{d}] \times [0,1]$. Total VRE output is then given by wx, where x is the capacity of VRE, which has capital cost per unit denoted by k_w. We denote the marginal density of w by $\mu(w)$ and its marginal distribution by $M(w)$, where $\mu(w)$ is assumed to be positive, continuous and bounded on $(0,1)$. We set $\mu(w) = 0$ outside that range. Hence, $M(w) = 0$ for $w \leq 0$ and $M(w) = 1$ for $w \geq 1$. The marginal cost of VRE is set to zero.

Demand net of VRE output, $d - xw$, is a random variable with the continuous probability distribution $G(x,y)$, which can be determined from

$$G(x,y) = \Pr(d - xw \leq y) = \int_0^1 \int_0^{y+xw} f(z,w)dzdw. \tag{2}$$

Note that this equation is also applicable when $y \leq 0$, because by assumption $f(z,w) = 0$ when $z \leq 0$. Hence, the inner integral is zero whenever $y + xw \leq 0$. We assume that excess supply can be curtailed at no cost, so the price is zero whenever net demand is negative. The density of net demand is $g(x,y) = \frac{\partial}{\partial y}G(x,y)$, which is continuous in x and y since

$$g(x,y) = \int_0^1 f(y + xw, w)dw.$$

Observe that

$$\frac{\partial}{\partial x}G(x,y) = \int_0^1 wf(y + xw, w)dw \geq 0.$$

The total production cost of meeting demand q from conventional generation is

$$C(q) = \int_0^q c(z)dz.$$

The maximum thermal electricity supply is $q(\overline{c})$. Net demand exceeds this with probability $1 - G(x, q(\overline{c}))$, in which case demand is rationed. The expected thermal operating cost of a curve $q(c)$ is therefore

$$P(x) = \int_0^{q(\overline{c})} g(x,y)C(y)\,dy + C(q(\overline{c}))\left(1 - G(x, q(\overline{c}))\right). \tag{3}$$

Operating costs are zero whenever net demand is negative. Thus, the lower integration limit can be set to 0. The first term integrates over all realisations where net demand is less than total thermal capacity; the second covers rationing events. The total investment cost of thermal power is:[19]

$$K(x) = q(\underline{c})k\left(\underline{c}\right) + \int_{\underline{c}}^{\overline{c}} k(c)q'(c)dc.$$

The expected rationed energy is

$$R(x) = \int_{q(\overline{c})}^{\overline{d}} g(x,y)(y - q(\overline{c}))dy. \tag{4}$$

We denote served demand by $\widetilde{d}$, which is smaller than d when there is loss of load.

We let $e\left(c\right)$ be the CO_2 emissions per unit of energy produced for a thermal plant with marginal cost c. We assume that $e\left(c\right)$ is continuous and bounded. Emissions may not be monotonic in c. For example, some natural gas plants may have lower emissions but higher fuel costs than some coal plants. Nevertheless, we assume that emissions from a unit with zero marginal cost are zero and that emissions are strictly positive for a unit with a positive marginal cost. The total emissions for a power system with VRE capacity x are denoted by $E\left(x\right)$.

3 Analysis

In Stage 2, thermal plants are dispatched in merit order, to minimize operating costs. This means that thermal units with a low marginal cost have a high probability of being dispatched, whereas units with a high marginal cost have a low probability of being dispatched. The maximum dispatch probability of a thermal unit is equal to $1 - G\left(x, 0\right)$, i.e. the probability that demand net of VRE output is positive. Each unit's dispatch probability is determined by its position on the supply curve. The following lemma identifies, for each dispatch probability, the unique thermal technology that minimizes the total expected cost of investment and operation.

Lemma 1 *The baseload technology with $c = 0$ is the uniquely optimal thermal technology for units with high dispatch probabilities, weakly above $-k'(0)$. For lower dispatch probabilities in the range $(0, -k'(0))$, the uniquely optimal thermal technology is implicitly determined by $-k'(c) = \pi\left(c\right)$.*

[19] We have added the term $q(\underline{c})k\left(\underline{c}\right)$, which is relevant if $q(\underline{c}) > 0$. This can, for example, occur for baseload when $\underline{c} = 0$. Holmberg and Ritz (2021) did not consider this term, which may have been an omission, but it does not influence the optimality conditions.

Not all thermal technologies are worth investing in. Technologies with very high marginal costs are dominated by demand rationing. From a welfare perspective, it would be more efficient to compensate consumers at the reservation price V than to build such plants. The next lemma shows, similar to Holmberg and Ritz (2021), that there is a sharp upper cutoff $\overline{c}$, determined solely by the investment-cost function $k(c)$ and the reservation price V. There are no investments in thermal power with a marginal cost above $\overline{c}$.

Lemma 2 *The upper cutoff* $\overline{c} \in (0, V)$ *can be uniquely determined from:*

$$k(\overline{c}) + k'(\overline{c})(V - \overline{c}) = 0. \tag{5}$$

There are no investments in thermal technologies with marginal costs $c > \overline{c}$. *Each thermal technology with* $c \in (0, \overline{c})$ *has one dispatch probability* $-k'(c)$, *for which it is more efficient than demand rationing and other thermal technologies. Similarly, the baseload technology with* $c = 0$ *is more efficient than demand rationing and other thermal technologies when the dispatch probability satisfies* $\pi \geq -k'(0)$.

When VRE capacity is sufficiently high, the most fuel-efficient technologies may also become uneconomic, because they require a high dispatch probability that can no longer be sustained. The following lemma characterises this lower cutoff $\underline{c}$.

Lemma 3 *There are no investments in thermal technologies unless net demand is positive with a positive probability, i.e.* $1 - G(x, 0) > 0$. *In this case, the lower cutoff* $\underline{c}$ *can be uniquely determined from:*

$$\begin{cases} -k'(\underline{c}) = 1 - G(x, 0) & \text{if } 1 - G(x, 0) < -k'(0) \\ \underline{c} = 0 & \text{if } 1 - G(x, 0) \geq -k'(0). \end{cases} \tag{6}$$

If $1 - G(x, 0) < -k'(0)$, *then* $\underline{c} > 0$ *and there are no investments in thermal technologies with marginal cost* $c < \underline{c}$. *If* $1 - G(x, 0) > -k'(\overline{c})$, *then* $\underline{c} < \overline{c}$, *and there will be investments in all thermal technologies with marginal cost* $c \in (\underline{c}, \overline{c})$.

It follows from the lemmas above that a thermal technology with $c \in (\underline{c}, \overline{c})$ is an optimal investment for a specific dispatch probability $\pi(c) = -k'(c)$, and no other dispatch probability. When $\underline{c} = 0$, the baseload technology with $c = 0$ is an optimal investment for dispatch probabilities weakly greater than $-k'(0)$. Given investments made, it will, in Stage 2, be cost effective to minimize the total operating costs. Hence, a unit with a high operating cost should not be running, unless every unit with a lower marginal cost is also running. Thus, the dispatch probability π of a unit with marginal cost $c \in (\underline{c}(x), \overline{c}]$ is given by $1 - G(x, q(c))$, the probability that net demand exceeds $q(c)$. This is also true for the last dispatched unit with $c = 0$ if $1 - G(x, 0) \geq -k'(0)$. Otherwise, we have $-k'(\underline{c}) = 1 - G(x, 0)$, which implies that $q(\underline{c}) = 0$. This can be summarised as follows.

Corollary 1 *The optimal supply curve of thermal power can be implicitly determined from:*

$$1 - G(x, q(c; x)) = -k'(c) \tag{7}$$

for $c \in (\underline{c}(x), \overline{c}]$.[20] *Similarly, we have* $1 - G(x, q(0)) = -k'(0)$ *if* $1 - G(x, 0) \geq -k'(0)$. *We have* $q(c) = 0$ *for* $c \leq \underline{c}$ *if* $1 - G(x, 0) < -k'(0)$.

In particular, we have

$$1 - G(x, q(\overline{c}; x)) = -k'(\overline{c}). \tag{8}$$

It follows from (5) that $\overline{c}$ does not depend on x. Hence, the probability that net demand exceeds $q(\overline{c})$, so that rationing is needed, does not depend on x. The loss of load probability is independent of x. This result breaks down for extremely large VRE investments for which net demand is below zero with a sufficiently high probability, so that $\underline{c} > \overline{c}$, i.e. all thermal technologies are driven out of the market. Beyond this point, the LOLP will begin to decline. It follows from implicit differentiation of (7) and our assumptions that $q'(c; x)$ is well defined for $(\underline{c}(x), \overline{c})$.

The optimality conditions in Lemmas 1-3 and Corollary 1 characterise the cost-minimizing thermal supply curve. The following proposition shows that this cost optimum is also a competitive market equilibrium with free entry, in which every active technology earns zero profit.

Proposition 1 *The optimality conditions in Lemmas 1-3 and Corollary 1 give a market equilibrium with zero expected profit for all thermal technologies* $c \in [\underline{c}(x), \overline{c}]$. *Any investments in technologies outside this range yield negative expected profit.*

We now examine how VRE entry affects thermal investments and the technology cutoffs. The next lemma shows that VRE entry uniformly reduces the thermal capacity for each marginal cost $c \in [\underline{c}(x), \overline{c}]$, which is related to Holmberg and Ritz (2021). In particular, this means that VRE entry reduces

[20] This condition is similar to Holmberg and Ritz (2021). The difference is that we allow for $\underline{c} > 0$, which is relevant for high penetration of VRE.

the capacity of baseload with zero marginal cost (nuclear and geothermal). When VRE penetration is high, it will push the lower technology cutoff upward, driving out the most fuel-efficient plants. The upper cutoff, by contrast, is invariant to VRE entry.

Lemma 4 *VRE entry reduces thermal supply* $\frac{\partial q(c;x)}{\partial x} < 0$ *for* $c \in (\underline{c}(x), \overline{c}]$*. If the probability that net demand is positive is sufficiently low,* $1 - G(x, 0) < -k'(0)$*, then VRE entry increases the lower technology cutoff,* $\frac{\partial \underline{c}}{\partial x} > 0$*. The upper technology cutoff* $\overline{c}$ *is not influenced by VRE entry, as long as* $\underline{c} < \overline{c}$*.*

In a long-run equilibrium, it will always be optimal to dispatch a technology $c \in (\underline{c}(x), \overline{c})$ with probability $-k'(c)$. This probability is the same, irrespective of VRE entry x. In a perfectly competitive market, such a plant will be dispatched whenever the market price is above c. Thus, we obtain the following result for the price distribution.

Corollary 2 *The probability distribution of prices remains the same above* $\underline{c}(x)$*, irrespective of VRE entry* x*. For high VRE entry, we have* $\underline{c}(x) > 0$*. If the price is below* $\underline{c}(x)$ *then the price is zero with certainty.*

It follows from Corollary 2 that, in a long-run equilibrium, there is no merit-order effect for moderate entry of VRE, where $1 - G(x, 0) > -k'(0)$, so that $\underline{c} = 0$. For high VRE entry, we have $\underline{c}(x) > 0$. The probability that the price equals zero is given by $G(x, q(\underline{c}(x)))$. The price is never in the range $(0, \underline{c}(x))$.

Corollary 3 *Prices for lower levels of VRE entry will first-order stochastically dominate prices for higher VRE levels. Irrespective of their risk preferences, consumers with flat demand will strictly prefer prices in a system with high levels of VRE entry such that* $1 - G(x, 0) < -k'(0)$*.*

One implication is that there is a long-run merit-order effect for high penetration of VRE.

The *conditional hazard rate* of demand given w is defined to be

$$\lambda(d, w) = \frac{f(d, w)}{\int_d^{\overline{d}} f(z, w)\, dz}.$$

It follows from Lemma 2 and (8) that the (unconditional) LOLP is independent of VRE entry, provided that some thermal capacity remains in the market. However, the volume of rationed energy may still change as x increases. To see why, consider the LOLP conditional on the VRE state, which does depend on w. If the conditional hazard rate $\lambda(d,w)$ is increasing in d (and non-decreasing in w), the demand distribution has relatively thin tails. Hence, once demand is sufficiently high to cause rationing at a given VRE output w, it is unlikely to exceed the scarcity threshold by much. As a result, the conditional LOLP declines rapidly as w increases, because higher VRE output shifts the threshold further into the tail. Consequently, rationing occurs primarily at low w, and high-w realisations contribute only modestly to reducing the expected volume of rationed energy. This leads to the following result.

Proposition 2 *Suppose $\underline{c} < \bar{c}$ and that the conditional hazard rate $\lambda(d,w)$ satisfies $\frac{\partial}{\partial d}\lambda(d,w) > 0$ and $\frac{\partial}{\partial w}\lambda(d,w) \geq 0$. In this case, the expected rationed energy is increasing in x.*

The volume of rationed energy affects consumer expenditure, since non-rationed demand is priced at the reservation price V when rationing occurs. Building on Proposition 2, the following result identifies conditions under which additional VRE nonetheless reduces total consumer expenditure.

Proposition 3 *Suppose $\underline{c} < \bar{c}$ and that the conditional hazard rate $\lambda(d,w)$ of demand satisfies $\frac{\partial}{\partial d}\lambda(d,w) > 0$ and $\frac{\partial}{\partial w}\lambda(d,w) \geq 0$. In this case, the expected total consumer expenditure when demand is stochastic is strictly decreasing in positive x.*

When the hazard-rate conditions of Proposition 3 are not satisfied, there are circumstances in which increased investment in renewable generation may lead to higher total expenditure in the long run. This can arise when demand exhibits long tails and renewable output is correlated with demand, so that additional renewable capacity reduces rationed volumes even when conventional capacity adjusts in a way that leaves the loss of load probability (LOLP) unchanged. In such cases, a larger volume of energy is delivered to consumers, which may increase their total expenditure—an outcome that need not be problematic from the consumer's perspective.

Consumer expenditure is given by $E[\widetilde{d}\,p]$, where $\widetilde{d}$ denotes served demand, which is below demand d when rationing occurs. As the next proposition shows, the result is cleaner for $E[d\,p]$, which disregards rationing. This would, for example, be the case if consumers had an unbounded willingness to pay for electricity, so that the reservation price is infinite.

Proposition 4 *If $\underline{c} < \bar{c}$ and consumers have uncertain demand, then $\mathbb{E}[d\,p]$ is strictly decreasing in positive x. Moreover, the covariance between price and demand is also strictly decreasing, as long as $\underline{c} = 0$, i.e. as long as the VRE penetration is moderate.*

It follows from Corollary 2 that there is no merit-order effect for moderate VRE entry. Still, there is a cannibalization effect, because there is negative correlation between the output of VRE and the price, and this effect strengthens when more VRE enters the market.

Proposition 5 *If $\underline{c} < \bar{c}$, the expected revenue of a VRE plant, $\mathbb{E}[\,w\,p\,]$, is continuous and strictly decreasing in the VRE capacity x. Moreover, assuming $\underline{c} = 0$ (moderate VRE penetration), the covariance between VRE output and price,* $\mathrm{Cov}(w,p) = \mathbb{E}[wp] - \mathbb{E}[w]\,\mathbb{E}[p]$, *is also strictly decreasing in x.*

Proposition 5 implies that VRE investments become less profitable when x increases. It follows from our assumptions that w is positive with probability 1, irrespective of d. Hence, sufficiently large VRE entry makes the curtailment probability ρ_c arbitrarily close to one. Since $(1-\rho_c)\,V$ is an upper bound on VRE revenue per unit of capacity, expected VRE revenue eventually falls below k_w. In our analysis, we have derived global optimality conditions for thermal power, which determine $c(y;\,x)$ and $q(\bar{c};x)$, for any level of VRE investment x and thermal output y. Thus, we can conclude the following from (26).

Corollary 4 *VRE investments are globally optimal when*

$$\mathbb{E}[w\,p] = \int_0^1 \left[\int_{z=q(0;x)+wx}^{q(\bar{c})+wx} w\,c(z-wx;\,x)\,f(z,w)\,dz + \int_{z=q(\bar{c})+wx}^{\bar{d}} w\,V\,f(z,w)\,dz \right] dw = k_w.$$

Such an optimum exists, and is unique, if $\mathbb{E}[w\,p] > k_w$ when $x = 0$.

The government could subsidise the investment cost, which would reduce k_w, or pay some sort of feed-in premium, which would be added to the revenue $\mathbb{E}[w\,p]$. We have shown that $\mathbb{E}[w\,p]$ is strictly decreasing with respect to VRE entry x, so such policy interventions would increase VRE entry in our model, if we disregard that market participants might have to pay taxes or fees to cover the cost of such interventions. Similarly, technological development that reduces the investment cost of VRE would also increase VRE entry.

4 Constant demand

In this section, we consider the special case of fixed demand $d = d^*$, which is interesting for two reasons. First, we will get explicit expressions for the optimal power system. Second, the effect of VRE on thermal operating costs is particularly stark, as flexibility is only needed to balance VRE variability.

Recall that $\mu(w)$ and $M(w)$ are the density and distribution of w. From (2) we obtain the net-demand distribution:

$$G(x,y) = \Pr(d^* - xw \leq y) = \int_{(d^*-y)/x}^{1} \mu(w)dw = 1 - M\left(\left(d^* - y\right)/x\right) \quad (9)$$

with density

$$g(x,y) = G_y(x,y) = \mu(\left(d^* - y\right)/x)/x \quad (10)$$

if $x > 0$.

Proposition 6 *Assume that demand is constant at d^* and that $1-G(x,0) > -k'(\overline{c})$, so that $\underline{c} < \overline{c}$. In this case, the optimal supply curve for $c \in (\underline{c}(x), \overline{c})$ is given by:*

$$q(c) = d^* - x\widehat{w}(c).$$

where

$$\widehat{w}(c) = M^{-1}(-k'(c)) \quad (11)$$

is the quantile of w for the dispatch probability $-k'(c)$. The high-entry threshold is at

$$x^* = \frac{d^*}{\widehat{w}(0)}.$$

Above the high-entry level the inverse relationship between $\underline{c}$ and x is given by:

$$x = \frac{d^*}{\widehat{w}(\underline{c})}. \quad (12)$$

Before VRE enters the market, all thermal capacity is zero-marginal-cost baseload, so the supply is d^* for all positive prices. Thus, the total fuel cost is zero.[21] As x increases, thermal supply will decrease linearly with the proportionality factor $\widehat{w}(c)$ at each price $c \in (\underline{c}, \overline{c})$. The functions M, M^{-1} and $-k'(c)$ are monotonic. Hence, supply will decrease by a larger amount at low $c \in (\underline{c}, \overline{c})$, so that $q(c)$ becomes a strictly increasing function for $c \in (\underline{c}, \overline{c})$ and $x > 0$. This means that the technology mix will also include mid-merit and peaker technologies, which have positive fuel costs.

The function $\widehat{w}(c) = M^{-1}(-k'(c))$ gives the renewable-state level $w = \widehat{w}(c)$ during hours when technology c is on the margin of being dispatched. To keep the dispatch probability of the technology unchanged when x increases, in accordance with Corollary 1, the thermal supply at c needs to shift inwards by $\widehat{w}(c)$.

[21] The price is indeterminate when both demand and supply are equal to d^*. To get a well-defined price, we can assume that there is a small volume of VRE capacity that shrinks to zero in the limit.

The function $\widehat{w}(c)$ has some policy relevance. For example, some electricity markets have a market-wide capacity mechanism that gives capacity an extra payment for being available, so that the loss of load probability can be reduced (Holmberg and Ritz, 2021). In such a case, each unit of available VRE capacity should receive a capacity payment scaled by the derating factor $\widehat{w}(\overline{c})$, which is the renewable-state level when the market price is $\overline{c}$, and the power system is on the margin of losing load.

In the constant-demand case, we can show that the expected rationed energy increases without imposing the hazard rate condition in Proposition 2.

Proposition 7 *Suppose demand is constant at* d^* *and* $1-G(x,0) > -k'(\overline{c})$, *so that there are some investments in thermal power. Then the expected rationed energy is linear in* x.

$$R(x) = x\int_0^{M^{-1}(-k'(\overline{c}))} \mu(z)(M^{-1}(-k'(\overline{c})) - z)dz.$$

The result for the rationed energy is consistent with the results for thermal technologies in Proposition 6. For each technology $c \in (\underline{c}(x), \overline{c})$, the capacity density $q'(c) = -x\widehat{w}'(c)$ and supplied energy density are proportional to x. At the same time, the dispatch probability $\pi(c) = -k'(c)$ is independent of x.

A central question is how VRE entry affects total fuel costs when demand is constant, so that all flexibility needs are driven by VRE variability. The following proposition shows that operating costs rise linearly with moderate VRE entry. Total fuel costs can begin to decline once entry is sufficiently far beyond the high-entry threshold, where baseload technologies have fully exited the market.

Proposition 8 *For constant demand, total expected operating costs are given by*

$$P(x) = x\int_{\underline{c}(x)}^{\overline{c}} ck'(c)\widehat{w}'(c)\,dc \quad (13)$$

$$\widehat{w}'(c) = -\frac{k''(c)}{\mu(M^{-1}(-k'(c)))}. \quad (14)$$

$P(x)$ *is a linear function for moderate entry,* $x \leq x^*$. *For* x *above the high-entry threshold* x^* *we get:*

$$P'(x) = \int_{\underline{c}(x)}^{\overline{c}} ck'(c)\widehat{w}'(c)\,dc + \underline{c}k'(\underline{c})\widehat{w}(\underline{c}), \quad (15)$$

which is negative for sufficiently large x.

Increasing operating costs are driven by the function $\widehat{w}'(c)$ which shows the extent to which a technology c is replaced by a less fuel-efficient technology $c + \Delta c$ when VRE enters the system. It follows from (13) that such replacements have little effect on the operating costs when c is low (baseload) and for peaker technologies with a low dispatch probability $-k'(c)$. The effect on total operating costs is largest for mid-merit technologies where $-ck'(c)$ is relatively high. Thus, it follows from (14) that the VRE-driven increase in total operating costs is largest when $k(c)$ is highly convex for mid-merit technologies and when renewable outcomes in the mid range are relatively infrequent. In such cases a small change in c has a large impact on the dispatch probability $-k'(c)$ and the associated renewable outcome. On the other hand, the increase in operating costs is less severe when $k(c)$ is approximately linear for mid-merit technologies and when renewable outcomes in the mid range are relatively frequent.

Above the high-entry threshold x^*, baseload technologies $c = 0$ will completely exit the market. Beyond this point, further VRE starts to reduce the output of technologies with $c > 0$, which will eventually reduce fuel costs. This explains the second term in (15). The second term will dominate earlier for large $\widehat{w}(\underline{c})$, which gives large shifts in the supply curve at low prices and early exit of baseload. This is the case when baseload technologies $c = 0$ need a dispatch probability close to 1 in order to be cost effective and when the renewable-state level $w = \widehat{w}(0)$ is close to 1 during hours when technology $c = 0$ is on the margin of being dispatched.

We have assumed that $e(c) > 0$ for $c > 0$. Hence, an argument similar to the proof of Proposition 8 gives the following result for emissions.

Proposition 9 *For constant demand, total expected emissions are given by*

$$E(x) = x \int_{\underline{c}(x)}^{\overline{c}} e(c)\, k'(c) \widehat{w}'(c)\, dc \tag{16}$$

$$\widehat{w}'(c) = -\frac{k''(c)}{\mu(M^{-1}(-k'(c)))}, \tag{17}$$

where $E(x)$ is a linearly increasing function for moderate entry, $x \leq x^*$. For x above the high-entry threshold x^* we get:

$$E'(x) = \int_{\underline{c}(x)}^{\overline{c}} e(c)\, k'(c) \widehat{w}'(c)\, dc + e(\underline{c})\, k'(\underline{c}) \widehat{w}(\underline{c}), \tag{18}$$

which is negative for sufficiently large x.

Thus, if demand is constant, then expected emissions will increase linearly with VRE entry as long as entry is moderate, and expected emissions will start to decrease for sufficiently high penetration. This suggests that reduced investment costs for VRE may not reduce emissions automatically, and that a cap on emissions may also be needed.

Based on our assumptions and results, we would argue that the risk of increased CO_2 emissions is highest in jurisdictions that lack both carbon pricing and hydropower, and that combine a large share of fossil-fuel generation with fuel-efficient baseload technologies such as nuclear power or geothermal energy. Examples of such jurisdictions are: Belarus and the United Arab Emirates as well as the U.S. states Georgia, North Carolina, and South Carolina. In each of these jurisdictions, at least 19.6% of electricity production comes from nuclear power, at least 39% from fossil fuels and at most 3.3% from hydropower.[22] Nevada, which produces electricity from 10% geothermal energy, 63% fossil-based fuels and 3% hydropower, is also a potentially problematic case. Belgium, Czechia and Hungary are three EU countries with a high share of nuclear power and fossil-fuelled production and only small amounts of hydropower. There is a risk that national emissions might increase in these three countries due to VRE entry, even if the EU-wide cap would prevent emissions from increasing in the EU as a whole.

5 Conclusions

We have analysed a long-run competitive equilibrium with free entry and a continuum of thermal technologies, and studied how the entry of variable renewable electricity (VRE) reshapes investments, prices and emissions. This is an equilibrium, in which the capacities of all technologies have fully adjusted to the entry of VRE. Our model neglects hydropower and domestic carbon caps and assumes that the market has fuel-efficient baseload technologies such as nuclear power or geothermal energy. Relevant examples include Belarus, Belgium, Czechia, Hungary, the United Arab Emirates, as well as the U.S. states Georgia, Nevada, North Carolina, and South Carolina.

We find that there is a lower cutoff $\underline{c}$ such that all technologies with marginal cost below $\underline{c}$ exit the market for high VRE entry. Thermal capacity shifts toward less fuel-efficient technologies. For moderate VRE entry the cutoff is zero. In this case, the average price is unchanged and there is no long-run merit-order effect. However, $\underline{c}$ rises once VRE penetration passes the high-entry threshold, and then the average price will go down, also in the long run. A key finding is that VRE entry does not change the price

[22] Country-level data are from 2023 and have been collected from Our World in Data. State-level data are from 2025 and have been collected from the EPA (United States Environmental Protection Agency).

distribution above the lower cutoff $\underline{c}$ or the loss of load probability, but below $\underline{c}$ mass shifts from positive prices to zero.

The absence of a long-run merit-order effect implies that consumers with nearly flat demand, such as energy-intensive industries, do not benefit from moderate VRE entry. They need high VRE penetration to get a lower average price. For time-varying demand, expected expenditure is decreasing in VRE capacity (except in the case discussed below), because more VRE decreases the price–demand covariance. Thus, consumers with variable demand spend less on electricity even if the (unconditional) price distribution is unaffected.

We show that VRE entry does not influence the loss of load probability in the long-run, as long as some thermal capacity remains. Still, given that loss of load occurs, we identify circumstances where rationed energy will increase. A possible exception is when the demand distribution has heavy tails, so that VRE entry reduces expected rationed energy volumes, delivering more energy at the reservation price. In this case, VRE entry could potentially increase total expenditure.

On the producer side, the VRE-output-weighted price is decreasing in VRE entry, which corresponds to the cannibalization effect. Moreover, the covariance between VRE output and price decreases (becomes more negative) with more VRE entry, as long as it is moderate.

The shift towards less fuel-efficient technologies explains why, when total demand across all consumers is constant, total fuel use by thermal technologies and total emissions rise as VRE entry increases up to the high-entry threshold. The shift to less fuel-efficient technologies is particularly damaging for mid-merit technologies, which combine a relatively high dispatch probability with a high fuel cost. This shift is large when the capital cost as a function of the operating cost is highly convex for technologies in the mid range and when intermediate renewable output levels are relatively rare.

When demand is constant, baseload technologies need to exit the market before VRE entry will start to decrease total fuel costs and total emissions. This exit can occur for low VRE levels when baseload technologies need a high dispatch probability to be cost effective and when the VRE output is close to capacity during hours when baseload is on the margin of being dispatched. We believe that a domestic carbon cap would often be more effective than domestic renewable targets for reducing fuel consumption and CO_2 emissions, as a cap directly incentivises efficiency and low-carbon generation.[23]

Our results are less pronounced quantitatively for VRE entry that has a relatively high availability or when batteries that smooth out VRE produc-

[23] Ambec and Crampes (2019) show how various energy policies can be combined to compensate for a missing cap on CO_2 emissions.

tion would be part of each VRE unit. However, as long as the availability of VRE capacity is imperfect, results are qualitatively the same. In the long run, VRE entry will still increase expected CO_2 emissions in a power system with efficient baseload and constant demand, unless the system has a carbon cap, long-term energy storage or fossil-free thermal production.

References

[1] Ambec, S. and C. Crampes (2019). "Decarbonizing electricity generation with intermittent sources of energy." *Journal of the Association of Environmental and Resource Economists* 6(6), 1105-1134.

[2] Antweiler, W. and F. Muesgens (2021). "On the long-term merit order effect of renewable energies." *Energy Economics* 99, 105275.

[3] Biggar, D.R. and M. R. Hesamzadeh (2014). *The Economics of Electricity Markets*. IEEE Press.

[4] Boiteux, M. (1960). "Peak Load Pricing." *The Journal of Business* 33(2), 157-179.

[5] Borenstein, S. (2005). "The long-run efficiency of real-time electricity pricing." *The Energy Journal* 26(3), 93-116.

[6] Borenstein, S. (2012). "The private and public economics of renewable electricity generation." *Journal of Economic Perspectives* 26(1), 67-92.

[7] Brown, T., and L. Reichenberg (2021). "Decreasing market value of variable renewables can be avoided by policy action." *Energy Economics* 100, 105354.

[8] Bushnell, J., and K. Novan (2021). "Setting with the sun: The impacts of renewable energy on conventional generation." *Journal of the Association of Environmental and Resource Economists* 8(4), 759-796.

[9] Chao, H. P. (1983). "Peak Load Pricing and Capacity Planning with Demand and Supply Uncertainty." *Bell Journal of Economics* 14(1), 179-190.

[10] Chao, H. P. (2011). "Efficient pricing and investment in electricity markets with intermittent resources." *Energy Policy* 39(7), 3945-3953.

[11] Crew, M. and P. Kleindorfer (1986). *The Economics of Public Utility Regulation*. MIT Press.

[12] Cullen, J. A. and S. S. Reynolds (2023). "Market dynamics and investment in the electricity sector." *International Journal of Industrial Organization* 89, 102954.

[13] Eisenack, K. and M. Mier (2019). "Peak-load pricing with different types of dispatchability." *Journal of Regulatory Economics* 56(2), 105-124.

[14] Graf, C. and C. Marcantonini (2017). "Renewable energy and its impact on thermal generation." *Energy Economics* 66, 421-430.

[15] Green, R. and T. O. Léautier (2015). "Do costs fall faster than revenues? Dynamics of renewables entry into electricity markets." Working Paper TSE-591, Toulouse School of Economics.

[16] Green, R. and N. Vasilakos (2011). The long-term impact of wind power on electricity prices and generating capacity. University of Birmingham Economics Discussion Paper 11-09.

[17] Gutiérrez-Martín, F., R.A. Da Silva-Álvarez and P. Montoro-Pintado (2013). "Effects of wind intermittency on reduction of CO2 emissions: The case of the Spanish power system." *Energy* 61, 108-117.

[18] Hirth, L. (2013). "The market value of variable renewables: The effect of solar and wind power variability on their relative price." *Energy Economics* 38, 218-236.

[19] Hirth, L., F. Ueckerdt and O. Edenhofer (2015). "Integration costs revisited–An economic framework for wind and solar variability." *Renewable Energy* 74, 925-939.

[20] Holmberg, P. and R. Ritz (2021). "Optimal capacity mechanisms for competitive electricity markets." *The Energy Journal* 42(1), 1-34.

[21] Joskow, P. and J. Tirole (2007). "Reliability and competitive electricity markets." *The RAND Journal of Economics* 38(1), 60-84.

[22] Kaffine, D. T., B. J. McBee and S. J. Ericson (2020). "Intermittency and CO2 reductions from wind energy." *The Energy Journal* 41(5), 23-54.

[23] Ketterer, J. C. (2014). "The impact of wind power generation on the electricity price in Germany." *Energy Economics* 44(C), 270-280.

[24] Lamont, A. D. (2008). "Assessing the long-term system value of intermittent electric generation technologies." *Energy Economics* 30(3), 1208-1231.

[25] Léautier, T.-O. (2019). *Imperfect Markets and Imperfect Regulation: An Introduction to the Microeconomics and Political Economy of Power Markets.* MIT Press.

[26] Liski, M. and I. Vehviläinen (2020). "Gone with the wind? An empirical analysis of the equilibrium impact of renewable energy." *Journal of the Association of Environmental and Resource Economists* 7(5), 873-900.

[27] Newbery, D. (2016). The long-run equilibrium impact of intermittent renewables on wholesale electricity prices. Cambridge Working Paper in Economics 1601.

[28] Prol, J. L., K. W. Steininger and D. Zilberman (2020). "The cannibalization effect of wind and solar in the California wholesale electricity market." *Energy Economics* 85, 104552.

[29] Reichenberg, L., T. Ekholm and T. Boomsma (2023). "Revenue and risk of variable renewable electricity investment: The cannibalization effect under high market penetration." *Energy* 284, 128419.

[30] Renewable Energy Policy Network for the 21st Century (REN21) (2024). Renewables 2024: Global Status Report.

[31] Sensfuß, F., Ragwitz, M., and M. Genoese (2008). "The merit-order effect: A detailed analysis of the price effect of renewable electricity generation on spot market prices in Germany." *Energy Policy* 36(8), 3086-3094.

[32] Stoft, S. (2002). *Power System Economics: Designing Markets for Electricity.* IEEE Press.

[33] Turvey, R. (1968). "Peak-load pricing." *Journal of Political Economy* 76(1), 101-113.

[34] Valentino, L., V. Valenzuela, A. Botterud, Z. Zhou and G. Conzelmann (2012). "System-wide emissions implications of increased wind power penetration." *Environmental Science & Technology* 46(7), 4200-4206.

[35] Willems, B. and Y. Yu (2023). "Bidding and Investment in Wholesale Electricity Markets: Discriminatory versus Uniform-Price Auctions", TILEC Discussion Paper No. 2023-14, Tilburg University.

[36] Zöttl, G. (2010). "A Framework of Peak Load Pricing with Strategic Firms." *Operations Research* 58(6), 1637-1649.

Appendix

Proof. (Lemma 1) Assume that a unit is dispatched with probability $\pi \in (0,1]$. The expected total cost for such a unit with marginal cost c is:

$$\text{Total cost} = \Pi = k(c) + \pi c.$$

The optimal technology c minimizes Π subject to $c \geq 0$. Since,

$$\begin{aligned} \Pi' &= k'(c) + \pi \\ \Pi'' &= k''(c) > 0, \end{aligned}$$

the first-order condition $k'(c) + \pi = 0$ gives a global cost minimum among thermal technologies for $c > 0$. It follows from strict monotonicity of $k'(c)$ and $\lim_{c\to\infty} k'(c) = 0$ that a unique first-order solution exists when $\pi < -k'(0)$. For high dispatch probabilities $\pi \geq -k'(0)$, we get $c = 0$ as a boundary solution. This solution is a global cost minimum, among thermal technologies, as $\Pi' = k'(c) + \pi > k'(0) + \pi \geq 0$ for $c > 0$. ■

Proof. (Lemma 2) (1) implies that $k(c) - k'(c)(c - V) < 0$ for sufficiently small $c > 0$. We have $k(c) > 0$ and $k'(c) < 0$, which ensures that $k(c) - k'(c)(c - V) > 0$ for $c \geq V$, so (5) has no solutions in this range. It follows from the change of sign, continuity of $k(c)$ and $k'(c)$ ($k(c)$ is twice differentiable), and the intermediate value theorem that (5) has at least one solution in the range $(0, V)$. This solution must be unique, because

$$\frac{d}{dc}\left(k(c) - k'(c)(c - V)\right) = k'(c) - k'(c) - k''(c)(c - V) > 0 \tag{19}$$

for $c \in (0, V)$.

Next, we explore the implications of $\overline{c}$ for investments in thermal technologies. It follows from Lemma 1 that a thermal technology with marginal cost $c > 0$ can only be optimal for a dispatch probability equal to $-k'(c)$; it is dominated by other thermal technologies for other dispatch probabilities. The total cost when investing in a unit c for the dispatch probability $-k'(c)$ is:

$$\Pi = k(c) - k'(c)c.$$

An alternative is to instead curtail consumers with probability $-k'(c)$. The cost for this is:

$$-k'(c)V.$$

Subtracting this cost from Π yields:

$$k(c) - k'(c)(c - V).$$

It follows from the argument above that this difference is strictly negative for $c < \overline{c}$ and strictly positive for $c > \overline{c}$. Hence, for each thermal technology with marginal cost $c < \overline{c}$, the technology is more cost effective than other thermal technologies and demand rationing for the dispatch probability, $\pi = -k'(c)$. The same argument applies to the baseload technology with $c = 0$. For this technology, we also have $k(0) + \pi(0 - V) < k(0) - k'(0)(0 - V) < 0$ for $\pi > -k'(0)$. Thus, it follows from Lemma 1 that this technology is more cost effective than demand rationing and other thermal technologies for dispatch probabilities $\pi \geq -k'(0)$. But a technology with marginal cost $c > \overline{c}$ is dominated by demand rationing for the dispatch probability $-k'(c)$, and is dominated by other thermal technologies for lower or higher dispatch probabilities (see Lemma 1). ■

Proof. (Lemma 3) If $1 - G(x, 0) \leq 0$ then net demand is non-positive and so no investments in thermal power will be profitable. Hence, we assume that $1 - G(x, 0) > 0$. We have $\lim_{c \to \infty} k'(c) = 0$. Thus, there is some finite and sufficiently large c, such that $-k'(c) < 1 - G(x, 0)$. In case $-k'(0) > 1 - G(x, 0)$, it follows from the change of sign, continuity of $k'(c)$ ($k(c)$ is twice differentiable), and the intermediate value theorem that there exists a positive solution to the equation $-k'(\underline{c}) = 1 - G(x, 0)$. This solution must be unique since $-k'(c)$ is strictly decreasing. This also implies that $-k'(c) > 1 - G(x, 0)$ for $c < \underline{c}$. Thus, for thermal technologies $c < \underline{c}$, the dispatch probability $\pi(c) = -k'(c)$, for which the technology would be an optimal choice, is higher than the maximum dispatch probability. Hence, it follows from Lemma 1 that such technologies cannot be optimal. We have $-k'(c) < 1 - G(x, 0)$ for $c > \underline{c}$. For technologies $c \in (\underline{c}, \overline{c})$ some thermal unit will have the dispatch probability for which the technology is optimal among thermal technologies. This is true for all thermal technologies with $c < \overline{c}$ when $1 - G(x, 0) \geq -k'(0)$, because $-k'(0) \geq -k'(c)$. In this case, we can set $\underline{c} = 0$. We have $\underline{c} < \overline{c}$, if $1 - G(x, 0) > -k'(\overline{c})$. Otherwise, the maximum dispatch probability would be too low, so that demand rationing is more cost effective than any thermal technology.

■

Proof. (Proposition 1) It follows from Lemma 1 that technology $c \in (\underline{c}, \overline{c})$ is dispatched with probability $\pi(c) = -k'(c)$. The market is competitive, so the price is above c with the same probability. Thus, the probability density of the price is given by $-\pi'(c) = k''(c)$ for $c \in (\underline{c}, \overline{c})$. Net of fuel costs, the revenue from a unit with marginal cost $c \in [\underline{c}, \overline{c}]$ is:

$$\begin{aligned}
-\int_c^{\overline{c}} (a-c)\,\pi'(a)\,da + \pi(\overline{c})(V-c) &= \int_c^{\overline{c}} (a-c)\,k''(a)\,da - k'(\overline{c})(V-c) \\
&= [(a-c)\,k'(a)]_c^{\overline{c}} - \int_c^{\overline{c}} k'(a)\,da - k'(\overline{c})(V-c) \\
&= k(c) - k(\overline{c}) - k'(\overline{c})(V-\overline{c}) \\
&= k(c)\,,
\end{aligned}$$

where we have used integration by parts and (5). Hence, the net revenue is the same as the investment cost. An investment in a technology with $c < \underline{c}$ would be dispatched with the probability $\pi(\underline{c})$ and it would get the same gross revenue as the technology $\underline{c}$. However, for the dispatch probability $\pi(\underline{c})$, the total investment and operating costs will be higher for technology c, see Lemma 1, so its payoff will be smaller. Similarly, the payoff would be negative for investments in technologies $c > \overline{c}$. ■

Proof. (Lemma 4) Equations (7) and (8) are identities that are valid for any x. Thus, we can differentiate both sides with respect to x.

$$\begin{aligned}
-G_x(x, q(c;x)) - G_y(x, q(c;x))\frac{\partial q(c;x)}{\partial x} &= 0 \\
\frac{\partial q(c;x)}{\partial x} &= -G_x(x, q(c;x))/G_y(x, q(c;x)) < 0
\end{aligned}$$

for $c \in (\underline{c}(x), \overline{c}]$. If $1 - G(x,0) < -k'(0)$, it follows from Lemma 3 that

$$\begin{aligned}
-k''(\underline{c})\frac{\partial \underline{c}}{\partial x} &= -G_x(x,0) \\
\frac{\partial \underline{c}}{\partial x} &= G_x(x,0)/k''(\underline{c}) > 0.
\end{aligned}$$

It follows from (5) that the upper technology cutoff $\overline{c}$ is not influenced by VRE entry. ■

We now establish some technical lemmas that we will require for our main results.

Lemma 5 *For fixed x and $c \in (\underline{c}(x), \overline{c})$, $c(q;x)$ and $q(c;x)$ are inverse functions, so $c_x = -c_y q_x(c;x)$ and $q_c(c;x) = \frac{1}{c_y(q(c;x);x)}$.*

Proof. By definition $c(q;x)$ and $q(c;x)$ are inverse functions, so for any fixed $\hat{c} \in (\underline{c}(x), \overline{c})$, $c(q(\hat{c};x);x) = \hat{c}$. Differentiating with respect to x gives

$$c_x(q(\hat{c};x), x) + c_y(q(\hat{c};x);x)q_x(\hat{c};x) = 0.$$

Differentiating both sides with respect to $\hat{c}$ gives

$$c_y(q\left(\hat{c};x\right);x)q_c\left(\hat{c};x\right)=1.$$

■

Lemma 6 *If* $\chi\left(w\right)$ *is continuous, non-decreasing and such that* $\chi\left(0\right)<\chi\left(1\right)$*, then*

$$\int_0^1\chi\left(w\right)f(q\left(c;x\right)+xw,w)\left(q_x\left(c;x\right)+w\right)dw>0 \text{ for } c\in\left(\underline{c}\left(x\right),\overline{c}\right].$$

The inequality also applies to $c=0$ *if* $\underline{c}\left(x\right)=0$*.*

Proof. Differentiating the first-order conditions (7) and (8) with respect to x yields:

$$\int_0^1 f(q(c;x)+xw,w)(q_x(c;x)+w)dw=0 \text{ for } c\in\left(\underline{c}\left(x\right),\overline{c}\right], \tag{20}$$

where we have used the definition of $G\left(x,y\right)$ in (2). Note that $f(y+xw,w)=0$ if $y+xw>\overline{d}$. We realise from (20), properties of $\chi\left(w\right)$, negativity of $q_x\left(c;x\right)$, monotonicity of $\left(q_x\left(c;x\right)+w\right)$ with respect to w that there must be some w_0, such that $f(q(c;x)+xw,w)(q_x(c;x)+w)\geq 0$ for $w\geq w_0$ and $f(q(c;x)+xw,w)(q_x(c;x)+w)\leq 0$ otherwise. The weighting factor $\chi\left(w\right)$ puts strictly more weight on $w\geq w_0$. This, together with (20), gives the result. It follows from Corollary 1 that the first-order condition in (7) is also valid at $c=0$ if $\underline{c}\left(x\right)=0$. ■

An analogous argument yields the following companion lemma.

Lemma 7 *If* $\chi\left(w\right)$ *is continuous, non-increasing and such that* $\chi\left(0\right)>\chi\left(1\right)$*, then*

$$\int_0^1\chi\left(w\right)f(q\left(c;x\right)+xw,w)\left(q_x\left(c;x\right)+w\right)dw<0 \text{ for } c\in\left(\underline{c}\left(x\right),\overline{c}\right].$$

The inequality also applies to $c=0$ *if* $\underline{c}\left(x\right)=0$*.*

Lemma 8 *Suppose* $\frac{\partial}{\partial d}\lambda(d,w)>0$ *and* $\frac{\partial}{\partial w}\lambda(d,w)\geq 0$*. If* $\tau\left(w\right)$ *is an increasing function of* w *then* $\lambda(\tau(w),w)$ *is strictly increasing and* $\frac{1}{\lambda(\tau(w),w)}$ *is strictly decreasing.*

Proof. We have

$$\frac{d}{dw}\lambda(\tau(w), w) = \frac{\partial}{\partial d}\lambda(d,w)\tau'(w) + \frac{\partial}{\partial w}\lambda(d,w) > 0,$$

and $\frac{d}{dw}\left(\frac{1}{\lambda(\tau(w),w)}\right) = \frac{-1}{\lambda(d,w)^2}\frac{d}{dw}\lambda(\tau(w), w) < 0$. ■

Proof. (Proposition 2) The expected rationed energy is given by:

$$R(x) = \int_0^1 \left[\int_{q(\overline{c})+wx}^{\overline{d}} (z - q(\overline{c}) - wx)\ f(z,w)\, dz \right] dw.$$

Hence,

$$R'(x) = -\int_0^1 (\partial q(\overline{c};x)/\partial x + w)\ \left(\int_{q(\overline{c})+wx}^{\overline{d}} f(z,w)\, dz\right) dw.$$

This equation can be written as

$$R'(x) = -\int_0^1 (\partial q(\overline{c};x)/\partial x + w)\ \chi(w) f(q(\overline{c}) + wx, w) dw$$

where

$$\chi(w) = \begin{cases} \frac{\int_{q(\overline{c})+wx}^{\overline{d}} f(z,w)dz}{f(q(\overline{c})+wx,w)}, & \text{if} \quad q(\overline{c}) + wx \leq \overline{d} \\ 0, & \text{if} \quad q(\overline{c}) + wx > \overline{d}. \end{cases}$$

Since the conditional hazard rate of demand is increasing in d and non-decreasing in w, and $\tau(w) = q(\overline{c}) + wx$, Lemma 8 implies that $\chi(w)$ is decreasing in w. Thus Lemma 7 implies

$$R'(x) = -\int_0^1 (\partial q(\overline{c};x)/\partial x + w)\ \left(\int_{q(\overline{c})+wx}^{\overline{d}} f(z,w)\, dz\right) dw > 0. \qquad (21)$$

■

Proof. (Proposition 3) The expected expenditure of consumers is given by

$$\mathbb{E}[\widetilde{d}\,p] = \int_0^1 \left[\int_{q(0;x)+wx}^{q(\overline{c})+wx} z\, c(z{-}wx;\ x)\, f(z,w)\, dz + \int_{q(\overline{c})+wx}^{\overline{d}} (q(\overline{c}) + wx)\ V\ f(z,w)\, dz \right] dw, \qquad (22)$$

where the first (interior) term corresponds to $q(0;x) < y < q(\overline{c})$ with $p = c(y;x)$ and the second (cap) term to $y \geq q(\overline{c})$ with $p = V$. Note that the price is zero for $z < q(0;x) + wx$, which explains the lower integration limit of the first term. Note that $q\,(0;x)=0$ if $\underline{c}(x)>0$. In this case, the price is only zero when VRE is curtailed. Differentiate (22) using Leibniz's rule:

$$\frac{d}{dx}\,\mathbb{E}[\widetilde{dp}] = \int_0^1 \Bigg[\underbrace{\int_{q(0;x)+wx}^{q(\overline{c})+wx} z\,c_x(z-wx;\,x)\,f(z,w)dz}_{\text{(i) direct interior effect}}$$
$$- \underbrace{\int_{q(0;x)+wx}^{q(\overline{c})+wx} z\,w\,c_y(z-wx;\,x)\,f(z,w)\,dz}_{\text{(ii) interior shift in } y}$$
$$+ \underbrace{(q(\overline{c})+wx)\,\overline{c}f(q(\overline{c})+wx,w)\;(\partial q(\overline{c};x)/\partial x+w)}_{\text{(iii) upper boundary term 1}}$$
$$- \underbrace{(q(\overline{c})+wx)V f(q(\overline{c})+wx,w)\;(\partial q(\overline{c};x)/\partial x+w)}_{\text{(iv) lower boundary term 2}}$$
$$- \underbrace{(q\,(0;x)+wx)\,\underline{c}(x)\,f(q\,(0;x)+wx,w)\;(w+q_x\,(0;x))}_{\text{(v) lower boundary term 1}}$$
$$+ \underbrace{(\partial q(\overline{c};x)/\partial x+w)\;V\;\left(\int_{q(\overline{c})+wx}^{\overline{d}} f\,(z,w)\,dz\right)}_{\text{(vi) interior effect from term 2}}$$
$$\Bigg]\,dw. \tag{23}$$

The two first terms can be combined with Lemma 5 ($c_x=-c_y q_x$), giving

$$\int_{q(0;x)+wx}^{q(\overline{c})+wx} (z\,c_x(z-wx;\,x)\;-z\,w\,c_y(z-wx;\,x))\;f(z,w)\,dz$$
$$= \;-\int_{q(0;x)+wx}^{q(\overline{c})+wx} z\,(q_x\,(c\,(z-wx)\,;x)+w)c_y(z-wx;\,x)f(z,w)\,dz.$$

Given renewable state w, the integral is from demand $z=q\,(0;x)+wx$ to $z=\;q(\overline{c})+wx$. We let $z=q\,(c;x)+wx$, where c varies from $\underline{c}$ to $\overline{c}$.

$$dz=q_c\,(c;x)\,dc$$

This gives

$$
\begin{aligned}
& -\int_0^1 \int_{q(0;x)+wx}^{q(\overline{c})+wx} z\left(q_x\left(c\left(z-wx\right);x\right)+w\right)c_y(q\left(c;x\right);x)f(z,w)\,dzdw \\
= & -\int_0^1 \int_{\underline{c}}^{\overline{c}} \left(q\left(c;x\right)+wx\right)\left(q_x\left(c;x\right)+w\right)c_y(q\left(c;x\right);x)f(q\left(c;x\right)+wx,w)\,q_c\left(c;x\right)dcdw \\
= & -\int_{\underline{c}}^{\overline{c}} \left(\int_0^1 \left(q\left(c;x\right)+wx\right)\left(q_x\left(c;x\right)+w\right)f(q\left(c;x\right)+wx,w)\,dw\right) c_y(q\left(c;x\right);x)q_c\left(c;x\right)dc
\end{aligned}
$$

By Lemma 6 we have

$$
\int_0^1 \left(q\left(c;x\right)+wx\right)\left(q_x\left(c;x\right)+w\right)f(q\left(c;x\right)+wx,w)\,dw > 0,
$$

if $x > 0$. We also have that for fixed x, $q(c;x)$ and $c(q;x)$ are inverse functions, so $q_c\left(c;x\right) = \frac{1}{c_y(q(c;x);x)}$ for $c \in (\underline{c}, \overline{c})$, which gives

$$
-\int_0^1 \int_{q(0;x)+wx}^{q(\overline{c})+wx} z\left(q_x\left(c\left(z-wx\right);x\right)+w\right)c_y(q\left(c;x\right);x)f(z,w)\,dzdw < 0.
$$

By Lemma 6 we also have,

$$
\int_0^1 (q(\overline{c})+wx)\left(V-\overline{c}\right)f(q(\overline{c})+wx,w)\left(\partial q(\overline{c};x)/\partial x+w\right)dw > 0,
$$

so

$$
-\int_0^1 (q(\overline{c})+wx)\left(V-\overline{c}\right)f(q(\overline{c})+wx,w)\left(\partial q(\overline{c};x)/\partial x+w\right)dw < 0.
$$

The term $\left(q\left(0;x\right)+wx\right)\underline{c}(x)\,f(q\left(0;x\right)+wx,w)\ \left(w+q_x\left(0;x\right)\right)$ is zero if $\underline{c}(x) = 0$. If $\underline{c}(x) > 0$ then it follows from Corollary 1 that $q\left(0;x\right) \equiv 0$, so that $q_x\left(0;x\right) = 0$. Thus,

$$
-\int_0^1 \left(q\left(0;x\right)+wx\right)\underline{c}(x)\,f(q\left(0;x\right)+wx,w)\ \left(w+q_x\left(0;x\right)\right)dw \leq 0
$$

By the hazard-rate conditions assumed in this proposition, it follows from the inequality (21) in the proof of Proposition 2 that

$$
\int_0^1 \left(\partial q(\overline{c};x)/\partial x+w\right)\ V\left(\int_{q(\overline{c})+wx}^{\overline{d}} f\left(z,w\right)dz\right)dw < 0.
$$

Thus, we can conclude from (23) that: $\frac{d}{dx}\,\mathbb{E}[\widetilde{d}\,p] < 0$. ■

Proof. (**Proposition 4)** If we disregard that demand can be rationed, (22) can be written:

$$\mathbb{E}[d\,p] = \int_0^1 \left[\int_{q(0;x)+wx}^{q(\overline{c})+wx} z\,c(z{-}wx;\,x)\,f(z,w)\,dz + \int_{q(\overline{c})+wx}^{\overline{d}} z\,V\,f(z,w)\,dz \right] dw, \tag{24}$$

Differentiate (24) using Leibniz's rule:

$$\begin{aligned}\frac{d}{dx}\,\mathbb{E}[dp] = \int_0^1 \Bigg[& \underbrace{\int_{q(0;x)+wx}^{q(\overline{c})+wx} z\,c_x(z-wx;\,x)\,f(z,w)dz}_{\text{(i) direct interior effect}} \\ & - \underbrace{\int_{q(0;x)+wx}^{q(\overline{c})+wx} z\,w\,c_y(z-wx;\,x)\,f(z,w)\,dz}_{\text{(ii) interior shift in } y} \\ & + \underbrace{(q(\overline{c})+wx)\,\overline{c}f(q(\overline{c})+wx,w)\ (\partial q(\overline{c};x)/\partial x + w)}_{\text{(iii) upper boundary term 1}} \\ & - \underbrace{(q(\overline{c})+wx)V f(q(\overline{c})+wx,w)\ (\partial q(\overline{c};x)/\partial x + w)}_{\text{(iv) lower boundary term 2}} \\ & - \underbrace{(q\,(0;x)+wx)\,\underline{c}(x)\,f(q\,(0;x)+wx,w)\ (w+q_x\,(0;x))}_{\text{(v) lower boundary term 1}} \\ & \Bigg] dw. \end{aligned} \tag{25}$$

The terms are the same as in the proof of Proposition 3, except that now we do not have any interior effect from term 2. This simplifies our argument, as we do not need to consider the effect of VRE on the rationed energy or the conditional hazard rate conditions. Apart from this simplification, we can use the same argument as in the proof of Proposition 3 to show that $\frac{d}{dx}\,\mathbb{E}[d\,p] < 0$. The covariance of the price and demand is given by:

$$\text{Cov}\,(p,d) = \mathbb{E}\,[d\,p] - \mathbb{E}\,[d]\,\mathbb{E}\,[p]\,.$$

As long as $\underline{c} = 0$, it follows from Corollary 2 that the distribution of prices, including the mean price, does not change when x increases. Moreover, d is exogenous. Hence, Cov (p,d) decreases as long as $\underline{c} = 0$. ■

Proof. (Proposition 5) For a given realization of w and demand d, the net demand served by conventional generators is $y = d - wx$. The price equals

$c(y;x)$ (the marginal cost of supplying y units at VRE capacity x) when $0 < y < q(\overline{c})$, and equals the price cap V when $y \geq q(\overline{c})$ (i.e. net-demand exceeds the maximum conventional output $q(\overline{c})$). If demand is low so that $y = d - wx \leq q\,(0;x)$, then $p = 0$ and there is no contribution to the expected value of $w\,p$. Thus,

$$\mathbb{E}[w\,p] = \int_0^1 \left[\int_{z=q(0;x)+wx}^{q(\overline{c})+wx} w\,c(z-wx;\,x)\,f(z,w)\,dz + \int_{z=q(\overline{c})+wx}^{\overline{d}} w\,V\,f(z,w)\,dz \right] dw. \tag{26}$$

Differentiating with respect to x and applying Leibniz's rule gives

$$\begin{aligned}\frac{d}{dx}\,\mathbb{E}[w\,p] = \int_0^1 \Big[&\int_{z=q(0;x)+wx}^{q(\overline{c})+wx} w\,c_x(z-wx;\,x)\,f(z,w)\,dz \\ &- \int_{z=q(0;x)+wx}^{q(\overline{c})+wx} w^2\,c_y(z-wx;\,x)\,f(z,w)\,dz \\ &- w\,(V-\overline{c})\,f(q(\overline{c})+wx,w)\,(\partial q(\overline{c};x)/\partial x + w) \\ &- (w+q_x\,(0;x))\,w\,\underline{c}(x)\,f(q\,(0;x)+wx,w)\Big]\,dw. \end{aligned} \tag{27}$$

The derivative is well defined, which implies that $\mathbb{E}[w\,p]$ is continuous with respect to x. Using $c_x = -c_y\,q_x$ (see Lemma 5), the interior terms combine to

$$-\int_{z=q(0;x)+wx}^{q(\overline{c})+wx} w\,(q_x(c\,(z-wx)\,;x)+w)\,c_y(z-wx;x)\,f(z,w)\,dz. \tag{28}$$

Changing variables ($z = q(c;x)+wx$, $dz = q_c\,(c;x)\,dc$) and reordering integration gives

$$-\int_{\underline{c}(x)}^{\overline{c}} c_y(q\,(c;x)\,;x)\,q_c\,(c;x) \int_0^1 w\,(q_x(c;x)+w)\,f(q(c;x)+wx,w)dwdc. \tag{29}$$

We set $\chi\,(w) = w$, apply Lemma 6 and the relationship $q_c\,(c;x) = \frac{1}{c_y(q(c;x);x)}$ to show that this term is negative. Similarly, it can be shown that the upper (price-cap) term gives a nonpositive contribution.

The term $-\,(w+q_x\,(0;x))\,w\,\underline{c}(x)\,f(wx+q\,(0;x)\,,w)$ is zero if $\underline{c}(x) = 0$. If $\underline{c}(x) > 0$ then it follows from Corollary 1 that $q\,(0;x) \equiv 0$, so that $q_x\,(0;x) = 0$. Hence, the lower boundary term is nonpositive. Thus,

$$\frac{d}{dx}\,\mathbb{E}[w\,p] < 0. \tag{30}$$

A similar argument can be used for the special case when total demand is certain. The covariance result follows from the fact that $\mathbb{E}[w]$ is constant and that $\mathbb{E}[p]$ does not depend on x as long as $\underline{c}(x) = 0$ (see Corollary 2), so

$$\frac{d}{dx}\operatorname{Cov}(w,p) = \frac{d}{dx}\left(\mathbb{E}[wp] - \mathbb{E}[w]\mathbb{E}[p]\right) = \frac{d}{dx}\mathbb{E}[wp] < 0. \tag{31}$$

■

Proof. (Proposition 6) It follows from (9) that

$$\begin{aligned} 1 - G(x, q(c)) &= \Pr(d^* - xw \geq q(c)) \\ &= M(\frac{d^* - q(c)}{x}) \end{aligned}$$

for $c \in (\underline{c}, \overline{c})$. Thus we have from Corollary 1 that

$$-k'(c) = M(\frac{d^* - q(c)}{x}),$$

which gives

$$q(c) = d^* - xM^{-1}(-k'(c)).$$

VRE entry is moderate as long as $d^* > xM^{-1}(-k'(0))$. The high-entry threshold x^* occurs when $q\left(0; x\right) = 0$, giving

$$x^* = \frac{d^*}{M^{-1}(-k'(0))}.$$

Above that level the relationship between $\underline{c}$ and x can be implicitly determined from:

$$M\left(d^*/x\right) = -k'(\underline{c}).$$

■

Proof. (Proposition 7) It follows from (10), (4), and Proposition 6 that

$$R(x) = \int_{d^* - xM^{-1}(-k'(\overline{c}))}^{d^*} \mu(\left(d^* - y\right)/x)/x(y - d^* + xM^{-1}(-k'(\overline{c})))dy$$

Let $z = \frac{d^* - y}{x}$, so $-xdz = dy$.

$$\begin{aligned} R(x) &= \int_0^{M^{-1}(-k'(\overline{c}))} \mu(z)(-zx + xM^{-1}(-k'(\overline{c})))dz \\ &= x\int_0^{M^{-1}(-k'(\overline{c}))} \mu(z)(M^{-1}(-k'(\overline{c})) - z)dz. \end{aligned}$$

■

Proof. (Proposition 8) It follows from Proposition 6 that

$$q(c) = d^* - x\widehat{w}(c)$$

so

$$q'(c) = -x\widehat{w}'(c).$$

The expected thermal operating cost in (3) can be rewritten as follows:

$$\begin{aligned} P(x) &= \int_{\underline{c}(x)}^{\overline{c}} c\,(1 - G(x, q(c)))\,q'(c)dc \\ &= x\int_{\underline{c}(x)}^{\overline{c}} ck'(c)\widehat{w}'(c)\,dc, \end{aligned}$$

which shows that $P(x)$ is linear in x for $x \leq x^*$, where $\underline{c}(x) = 0$. We have from (11) that

$$\begin{aligned} \widehat{w}'(c) &= \left[\frac{d}{dz}\left(M^{-1}(z)\right)\right]_{-k'(c)}(-k''(c)) \\ &= -\frac{k''(c)}{\mu(M^{-1}(-k'(c)))} \\ &< 0 \end{aligned}$$

since k is strictly convex. Since $k'(c) < 0$ and $\widehat{w}'(c) < 0$, the integrand $ck'(c)\widehat{w}'(c)$ is positive, so

$$P'(x) > 0$$

for $x \leq x^*$. For x above the high-entry threshold we get from Leibniz's rule that

$$\begin{aligned} P'(x) &= \int_{\underline{c}(x)}^{\overline{c}} ck'(c)\widehat{w}'(c)\,dc - x\underline{c}'(x)\,\underline{c}k'(\underline{c})\widehat{w}'(\underline{c}) \\ &= \int_{\underline{c}(x)}^{\overline{c}} ck'(c)\widehat{w}'(c)\,dc + \underline{c}k'(\underline{c})\widehat{w}(\underline{c}), \end{aligned}$$

where we have used the identity that $x = \frac{d^*}{\widehat{w}(\underline{c})}$ for $x > x^*$, which follows from (12). The second term will eventually dominate for $x > x^*$. In particular, fuel costs will be zero when $\underline{c}(x) \to \overline{c}$, so that all thermal technologies completely exit the market. ■