WORKING PAPERS

N° 1555

July 2024

“Optimal Regulation of Electricity Provision with Rolling and Systemic Blackouts”

Catherine Bobtcheff, Philippe De Donder and François Salanié

Toulouse School of Economics

Optimal Regulation of Electricity Provision with Rolling and Systemic Blackouts*

Catherine Bobtcheff† Philippe De Donder‡ François Salanié§

July 21, 2024

Abstract

We set up a static model of electricity provision in which delivery to consumers is only imperfectly reliable. Blackouts can be either rolling or systemic; in both cases a price cap has to be imposed on the wholesale market. We characterize optimal allocations and we show that for any given value of the price cap on the wholesale market, one can decentralize these allocations thanks to two types of regulatory instruments: a retail tax, and capacity subsidies. Some properties follow. If demand is affected by multiplicative shocks only, capacity subsidies are exactly financed by the revenues from the retail tax. If moreover the distribution of systemic blackouts is exogenous, a price cap is sufficient, provided it is set at the value of lost load. In all other cases, all instruments are needed, and capacity subsidies need to be differentiated, based on the correlation between available capacity and its social value.

Keywords: electricity, reliability, renewables, climate change.

JEL Classification: D24, Q41, Q42, Q48.

---

*We thank Stefan Ambec, Claude Crampes, Shefali Khanna, participants to the 13$^{\text{th}}$ Toulouse Conference on the Economics of Energy and Climate, the 27$^{\text{th}}$ Annual Conference of the European Association of Environmental and Resource Economists (EAERE, Rimini), the X$^{\text{th}}$ Conference of the Spanish-Portuguese Association of Natural and Environmental Resources Economics (AERNA, Barcelona), and various experts from the energy sector, for their insightful suggestions and remarks. This research has benefitted from the financial support of Engie in the context of the research foundation TSE-Partnership. The authors gratefully acknowledge funding from the French National Research Agency (ANR) under its Programme d'Investissement d'Avenir ANR-17-EURE-0001 and ANR-17-EURE-0010.

†Paris School of Economics and CNRS. E-mail: catherine.bobtcheff@psemail.eu.

‡Toulouse School of Economics, University Toulouse Capitole and CNRS, Toulouse, France. E-mail: philippe.dedonder@tse-fr.eu.

§Toulouse School of Economics, University Toulouse Capitole and INRAE, Toulouse, France. E-mail: francois.salanie@tse-fr.eu.

1 Introduction

Climate change, and the policies undertaken to mitigate it, is deeply transforming the energy sector. The drive away from fossil fuels has led to the development of so-called intermittent sources that rely on sunlight or wind to produce electricity. Other energy sources such as hydrogen combustion also require electricity in the first place for the production of hydrogen. The consequence is that electricity will probably be needed for almost every use, from transportation to heating to cooling, and also for many industrial processes.

The provision of electricity will therefore require huge investments, both in production units and in the grid that delivers it to consumers. This paper argues that this costly transformation should also be analyzed in terms of risk and reliability. An electricity grid operates under strong technical constraints: in particular, supply and demand have to be precisely balanced in continuous time. This balance is difficult to manage when a significant part of supply is intermittent; when demand depends also on weather, and cannot be managed in real time; and when additional pressure is put on production units and on the grid because of more extreme climatic events. Reliability, including the possibility of major blackouts, is thus likely to become a major concern in the future, both for households and for industrial or commercial processes. In the absence of alternative energy sources, users will face a risk of disconnection. The February 2021 crisis in Texas illustrates perfectly the costs and hazards of the unreliability of electricity production and distribution.

This paper aims at analyzing these risks, and how they impact the optimal regulation of electricity markets. Our model of the electricity sector is inspired by Joskow and Tirole (2007). It is a static model that allows for heterogenous producers, incorporates intermittency and arbitrary shocks to costs, capacities and demand, and proposes an original modeling of disconnections, and of rolling and systemic blackouts. Our methodology is straightforward: We characterize optimal allocations (i.e., investments, and productions and consumptions in each state of nature), and we show how to decentralize these optima thanks to a regulatory intervention; in particular, we derive explicit formulas for the optimal values of regulatory instruments.

Let us begin by summarizing the main steps in our analysis. We emphasize that difficulties with the electricity market mainly originate in the demand side. At the most local level, agents consume electricity by connecting devices to their own private loop, which is itself connected to the grid. By switching on and off these devices, agents determine their electricity demand. For technical reasons, each individual demand is either fully satisfied, or not at all: partial rationing of a consumer's demand is so far impossible.[1] Another difficulty is behavioral: because attention is costly, the typical consumer is not able to react in real time to changes in the price he pays.[2]

These features severely constrain retail contracts. Strikingly, retailers all over the world have ended up selling the same type of contract: in exchange for a fixed fee, this contract allows consumers to demand the quantity they wish at a pre-specified retail price. In the following, we assume that all consumers sign this type of contract. This rigidity in the management of demand makes it impossible to balance the wholesale electricity market when demand lies above available capacity. In such a case, each producer becomes indispensable, making the surge in prices irresistible, while doing little to reduce the imbalance.

Regulators then have to step in, since the only way to balance the market is to ration demand. Interestingly, regulators typically choose to disconnect the minimum number of consumers consistent with the equality between demand and available capacity, and by doing so they perpetuate the market power of producers since all of them remain indispensable; we stick to this case in our model.[3] Moreover, in order to settle transactions between producers and retailers at a reasonable price, regulators typically impose a cap on the wholesale price. The value for this cap varies from regulator to regulator, and from country to country.[4]

[1]This would require setting priorities at the level of each device, or installing several distinct loops that could be switched on and off independently. In the future, one indeed expects an increase in the use of smart devices able to manage their activity as a function of an information on the real-time wholesale price of electricity. This article ignores this possibility.

[2]This is not the case for some industrial users. We choose to ignore this possibility in this article, for the sake of simplicity.

[3]Regulators display a willingness to avoid orderly blackouts which seems out of proportion with standard estimates of the cost of these episodes; on this point, see Wolak (2013) and Joskow (2022, pp. 5-6).

[4]Since 2022, the cap is set at €3,000 per MWh in Europe. The price cap in Texas was particularly high in 2021, at $9,000; it became binding in February 2021. This value was considerably higher than the

Unfortunately, setting a price cap creates two distortions, that can be corrected thanks to two new regulatory instruments. On the retail side, retailers benefit from a capped wholesale price of electricity, and therefore the competitive retail price is reduced: this may justify imposing a tax on retail electricity. On the supply side, the producers' revenues are reduced, and this makes investments in capacity less profitable. How to deal with this "missing money" problem is the topic of many works in the literature, and we simply deal with it by introducing the possibility of a subsidy to capacity.

In simple models, the missing money problem can be fully solved by setting the price cap at the level of the Value of Lost Load (VOLL), defined as the social cost of a power outage: indeed, such a level provides the right incentives to investors and producers. However, this intuitive argument raises a number of difficulties. The definition of the VOLL is not precise enough, as its value may for example depend on whether the outage is anticipated or not,[5] and its measurement raises numerous difficulties.[6] In our model, it is also dependent on contingencies such as temperature: a freezing cold makes heating indispensable. Finally, there are good reasons for reducing the price cap below the VOLL. With a high price cap, retailers are put into an uncomfortable situation: they sell electricity at a fixed retail price, while they buy it at a very volatile wholesale price. Even if risk-neutrality is assumed, when wholesale prices rise retailers may end up in a situation in which bankruptcy is the only option, as observed recently in the UK when gas prices surged. Setting a lower price cap is thus a manner to reduce the risk borne by retailers, and ultimately by consumers when these retailers go bankrupt. Moreover, it has long been argued that lowering the price cap reduces the exercise of market power by dominant producers. While these considerations are left unmodeled in this paper, they motivate our choice of taking the price cap as given, and deriving the associated optimal regulation.[7]

To summarize, our view is that the inability to flexibly manage demand in real time creates a market failure: it is sometimes impossible to balance the market. This in turn

average yearly price of $20 in 2020; still, the market could only be balanced thanks to demand rationing.

[5] See Joskow and Tirole (2007) for a precise discussion.

[6] See Gorman (2022).

[7] One might envision more complex interventions, such as making the price cap or the level of rationing contingent on market information ex-post, instead of being set ex-ante. Gerlagh, Liski and Vehviläinen (2022) allow for such possibilities in a stylized model of demand rationing with responsive and non-responsive consumers.

implies a variety of regulatory interventions: one has to ration demand, and to set a price cap. The price cap in turn requires to tax the retail price, and to subsidize capacity. But the story does not end here. States of the world where available capacity is lower than demand call for rationing demand through blackouts. In our model, two types of blackouts may occur. In a rolling (or orderly) blackout, all available capacity is allocated uniformly to consumers, who then face the same probability of not being served. In a systemic blackout, the allocation of the available capacity to consumers puts some additional stress on the transmission grid, triggering disconnections from the grid of some production units; such episodes can have dramatic consequences.

The point is that the types and sizes of blackouts are not only determined by exogenous shocks. They also depend on the existence of excess capacities that would allow to hedge against unanticipated surges in demand, or breakdowns in production, or incidents in the grid. We model this by allowing the impact of a blackout to depend on what we call the capacity index, namely the ratio of available capacity to aggregate demand. This is a plausible measure of the tension on the grid, which moreover neutralizes scale effects. A consequence is that every new investment in capacity, and every increase in the retail price, contributes to higher values of the capacity index. This in turn reduces both the likelihood and the size of the two types of blackouts. But this improvement in the performance of the grid is a public good likely to be under-provided in the absence of regulation.

Our first contribution is thus to set up a general model of the electricity sector, with heterogenous producers, arbitrary shocks on both supply and demand, and different types of blackouts. Efficiency is defined as the maximization of total surplus. Efficient investments, productions, and consumptions are characterized through simple first-order conditions. We propose an extended notion of competitive equilibria to study decentralization of the efficient allocations: firms are price-takers, unless demand is rationed, in which case they become indispensable and behave strategically. We finally derive the optimal regulation, for arbitrary values of the price cap.

We now state our results, and comment on their relationship to the literature. The main result characterizes the optimal regulation: we show that for any value of the exogenous price cap, if this cap is high enough to convince producers to use all available capacity, then one can decentralize the optimal allocations using only two types of instruments: a retail tax, and differentiated subsidies for capacity. Our model thus provides a general normative analysis for the electricity sector, which curiously was missing so far.

The model also emphasizes the symmetry between demand and supply: reducing demand is as important as fostering investments. This can be seen for the missing money problem: the price cap reduces the payments to producers, but it also reduces the costs borne by retailers when buying electricity on the wholesale market, and this distortion should also be corrected. Similarly, as soon as the size and occurrence of blackouts increase with demand and decrease with installed capacity, then capacity should be subsidized, and demand should be taxed. These simple arguments do not seem to appear in the literature, from Joskow and Tirole (2007) to Léautier (2016) to Elliott (2024), and neither in the survey of reliability by Borenstein, Bushnell and Mansur (2023).[8]

The optimal retail tax includes two expected terms, one for the missing money problem and one for the public good problem. Both terms are decreasing with the price cap: indeed, a higher cap makes electricity more expensive, which reduces demand, and increases the stability of the grid. It is then possible to pinpoint a value for the price cap such that the retail tax is nil; but we show that this value of the price cap is above the Value of Lost Load, in contradiction with the actual practice of regulatory bodies. This justifies the use of a retail tax, even in the absence of environmental issues such as global warming, and even in the absence of blackouts.

The optimal subsidies to capacity share a similar additive structure. Their values depend not only on the nominal capacity of the production unit under consideration, but also on how the effective capacity varies across states of nature. This effect is particularly important for intermittent sources: the subsidy to solar is reduced if solar panels produce electricity only when its social value is low. The precise formulas we provide for the optimal values of these regulatory instruments make it possible to evaluate this effect. We show that in the presence of intermittency, subsidies to capacity must be differentiated.

As we have said, the optimal value of these instruments depend not only on expected

[8] An exception is Borenstein and Holland (2005), discussing the choice between a retail tax and capacity payments. Their model has homogenous production plants, no intermittency and no blackouts, and this set of assumptions implies that one instrument only may be sufficient for efficiency.

values, but also on covariances between, say, the elasticity of demand and total consumption. These are complicated objects, and at this point we simplify the problem by assuming that demand is only subject to multiplicative shocks.[9] A striking property follows: at the optimum, the regulator's budget is balanced, i.e., the revenues from the retail tax exactly balance the subsidies to capacity. To the best of our knowledge, this result is new in the literature.[10] It may be valuable as a guideline for public agencies.

We can finally ask under which conditions a price cap is a sufficient instrument for optimality. We show that the missing money problem is solved if one sets the price cap equal to the expected value of the VOLL, conditional on demand being rationed; but this expectation is computed with nontrivial weights that depend inter alia on how the elasticity of demand varies with the state of nature. The formula simplifies somewhat miraculously in the case when demand is only affected by multiplicative shocks: then both the elasticity of demand and the VOLL turn out not to depend on the state of nature, so that one can set the price cap equal to the VOLL without ambiguity. But in general the other instruments still have a role to play: In fact, the only case in which a well-set price cap is sufficient for optimality is when demand is only affected by multiplicative shocks, and there is no intermittency, and blackouts are purely exogenous. This very restrictive set of conditions thus severely limits the interest of the so-called energy-only paradigm, which claims that a price cap equal to the VOLL is a sufficient instrument for optimality.

Our paper proposes an original modeling of both rolling and systemic blackouts, with the associated disconnections on both demand and supply, in a compact manner. A key role is played by a function which links the occurrence and size of blackouts to realized demand and available capacity, plus stochastic shocks. This function is in principle identified from existing data, though the econometrician would face the difficulty of establishing the probability of very rare events such as a large, systemic blackout. Both rolling and systemic blackouts are studied in Joskow and Tirole (2007), though in a simplified model,

[9]This powerful assumption is common in the literature; see for example Joskow and Tirole (2007, property (v) pp. 62), Elliott (2024, pp. 16), or Gowrisankaran, Reynolds and Samano (2016, pp. 1198). This last paper computes the value of intermittent electricity, using US data. It proceeds by solving directly the system operator problem of maximizing total surplus. More intermittency creates significant costs because storage is not available.

[10]The result appears in the 2003 working paper version of Borenstein and Holland (2005), in a model which is much simpler.

and without our optimality result for regulatory instruments. Llobet and Padilla (2018) study investment choices when blackouts foster a reputation loss. Elliott (2024) models only rolling blackouts, with disconnections of the supply side in proportion to the ratio of available capacity over demand.

Being essentially static, the model ignores important determinants of investments that are the focus of much more ambitious works. The fully dynamic model in Elliott (2024) incorporates dynamic investment strategies in a non-stationary environment: this allows to precisely study the timing of investments under various scenarios. It also allows for market power, both at the investment stages and on the wholesale market. A similar emphasis on investment timing can be found in Gowrisankaran, Langer and Zhang (2024). We also ignore the possibility of storages, using for example batteries, as studied in Butters, Dorsey and Gowrisankaran (2024). Overall, our model is much closer to the model in Joskow and Tirole (2007).

Another deliberate choice is to reduce the exercise of market power to the case when each producer becomes indispensable, i.e when demand is rationed because no price could possibly balance the market. Our position is that this non-existence is a characteristic of electricity markets which needs to be studied per se. It is an important source of market power that differs from the day-to-day, Cournot-style market power which admittedly exists on electricity markets, but is in no way different from market power on any other market. Such a stand is in opposition to much of the applied literature, that often aims at recovering information on costs from the observations of strategies. On the other hand, it allows us to provide a clear picture of market failures, and of the instruments needed to remedy these failures.[11] Finally, one may argue that market power matters essentially in times of crisis, when the social value of electricity can reach very high levels. In our model firms indeed become indispensable in case of demand rationing, and this is what drives the price upward, up to the price cap.[12]

[11]Keppler (2017) also argues in favor of identifying market failures in benchmark models: here, the public good problem of grid reliability, and the fact that some markets are missing that would allow customers to trade in real-time. Ferrasse, Neerunjun and Stahn (2022) also argue that power outages occur because markets are incomplete. Using a general equilibrium type of model, they show how endogenous retail contracts can remedy to this feature and eliminate both the non-existence problem and demand rationing. Assuming perfect competition is sometimes illuminating.

[12]Interestingly, firms with market power may behave so as to make such episodes more likely. McRae and Wolak (2019) claim to have identified such behaviors on the Columbian electricity market.

Finally, one main reason why intermittent sources of energy have been on the increase is that their use does not release carbon into the atmosphere. Introducing an optimal tax on emissions in our model would not change any of the results: it would simply modify the unit cost of different technologies. A sub-optimal tax creates more complex distortions, that could be corrected by subsidizing green electricity, and once more by creating a retail tax, as in Ambec and Crampes (2019). The model in Elliott (2024) has three market failures (pollution, inelastic final demand, market power) and two instruments (carbon tax and capacity payment). A carbon tax favors renewables, but makes blackouts more likely; capacity payments exert opposite effects. The model is estimated on Western Australia data in order to fund the right balance between these two instruments. Our view here is that optimality can be reached with a Pigovian carbon tax, and capacity payments and a retail tax set according to our formulas; but we ignore market power issues, as well as the dynamics of investment.

We proceed as follows. The next section presents the model. We then characterize optimal allocations in Section 3, competitive equilibria in Section 4, and a first-best regulation in Section 5. Section 6 concludes. Proofs and details of the derivations are gathered in the Appendix.

2 The Model

State of nature. We begin by defining an exogenous shock $s \in S$, with known distribution. This shock captures the effects on both supply and demand of weather, earthquakes, fires, breakdowns, and of all elements considered as exogenous– *i.e.*, beyond the control of agents.

Timing. Ex-ante, a regulatory policy is set, and electricity producers invest in production units, while retailers offer retail contracts to consumers. Then, the state of nature s is realized, and it is publicly observed by all agents. Ex-post, consumers choose their demand, producers decide what to produce, and the wholesale price is determined on the market, depending on stochastic blackouts. Retailers pay the wholesale price to producers and distribute electricity to consumers, or at least to those producers and consumers that are still connected to the grid.

Supply. Electricity is produced by different types of units, indexed by the subscript k. A unit of type k has a variable cost function $c_k^s(q_k)$ in state s, weakly convex in production $q_k \in [0, K_k^s]$. The upper bound is a finite production capacity, and we allow it to depend on the state s, so as to capture for example the dependence of intermittent production units on weather.

Ex-ante, it is possible to invest in an arbitrary number $x_k \geq 0$ of units of type k. For the sake of simplicity, we assume that the numbers x_k are strictly positive real numbers: this will avoid lengthy discussions of corner or integer solutions. The associated investment cost $I_k(x_k)$ is assumed weakly convex, so as to capture the idea that for some types of units, such as dams or wind farms, there is a limited number of possible locations with the same characteristics.

Given the vector of investments $X = (x_k)$, total investment cost is incurred ex-ante, before the realization of the state of nature, and it equals

$$\sum_k I_k(x_k).$$

After the realization of the state of nature s, total available capacity in state s is

$$K^s(X) \equiv \sum_k x_k K_k^s.$$

We also define the aggregate cost function in state s, given the investments X:

$$C^s(Q, X) = \min\{\sum_k x_k c_k^s(q_k) : 0 \leq q_k \leq K_k^s, \sum_k x_k q_k = Q\}.$$

We denote by C_Q^s the marginal production cost, and by AC^s the average cost. Finally, we let π_k be the profit function of a unit of type k:

$$\pi_k^s(p) = \max\{pq - c_k^s(q) : 0 \leq q \leq K_k^s\},$$

where p denotes the unit wholesale price of electricity. A useful identity is

$$\frac{\partial C^s}{\partial x_k}(Q, X) = -\pi_k^s(C_Q^s(Q, X)). \tag{1}$$

Demand. There is a mass of (possibly heterogenous) consumers, aggregated into a representative consumer, who derives a gross surplus $v^s(e)$ from the consumption of a quantity e of electricity in state s. The gross surplus function v^s is strictly concave in e, for each realization of the exogenous state of nature s.

The consumer gets electricity thanks to a retail contract. For technological reasons, it is impossible to partially disconnect a consumer: either a consumer sees his demand satisfied, or he is disconnected and gets nothing. Moreover, we assume it is too costly for the consumer to react in real time to new information about prices; while he may adapt his consumption in real time to different states of nature, such as changes in temperature. These elements create a strong rigidity in the management of demand. In particular, retail contracts simply specify a retail price $\overline{p}$, set before the realization of the state of nature s, and a fixed fee A. After s is publicly observed, demand $D^s(\overline{p})$ obtains from the maximization of the net surplus

$$v^s(e) - \overline{p}e.$$

This demand function is decreasing with the retail price $\overline{p}$, with a derivative $D_p^s < 0$. We also define the elasticity of demand

$$\varepsilon^s(p) = -\frac{pD_p^s(p)}{D^s(p)}. \tag{2}$$

Finally, when a consumer is disconnected, he obtains a gross surplus $v^s(0)$, and does not pay anything.

Now, suppose that global production e is reduced by one unit, so that one has to allocate this reduction across consumers. We assume that the only way is to disconnect some consumers, and that this disconnection is uniform across types of consumers. Hence, the regulator disconnects the same proportion y of each type, so that $y = 1/e$, and when disconnected the representative consumer experiences a loss $v^s(e) - v^s(0)$. The associated aggregate social loss is called the Value of Lost Load (VOLL), and it equals

$$\ell^s(e) = \frac{v^s(e) - v^s(0)}{e}.$$

Note that ℓ^s is decreasing with e, by concavity of the surplus function v^s. It represents the social loss associated to a reduction by one unit of electricity supply, when all consumers are connected and face the same retail price, and when our assumption of rigid demand management holds.[13] We shall assume that in any case ℓ is high enough to lie above any conceivable marginal cost of production.

[13]Note that, if demand could be flexibly managed, then the VOLL would be $v'(e)$, which is lower

Balancing demand and supply. The wholesale electricity market must be balanced at any time, and this may raise some difficulties, in particular when demand exceeds the available capacity, or when the grid is impaired, say by a storm or a fire. In such cases, the rigidity of demand imposes that some consumers be cut off from the grid. But disconnecting some consumers sometimes implies disconnecting simultaneously some of the production units. This also re-routs electricity towards transmission lines that may already be congested, possibly triggering additional disconnections. The outcome may then be either a rolling blackout (where some consumers are disconnected but all available capacity remains connected), or a systemic blackout (during which entire parts of the grid are out of function, so that some available capacity is off-line).[14]

To model this process in a simple but general manner, we assume that the type and the size of a blackout are a function of both the state of the world s, and of the following *capacity index*:

$$\kappa = \frac{K^s(X)}{D^s(\overline{p})}. \tag{3}$$

The capacity index κ measures the tension between nominal capacity and demand in state s. It presents the advantage of neutralizing scale effects. We moreover assume that disconnection occurs uniformly, both for production units and for consumers, so that disconnection affects only the size of each side of the market, and not its composition. As a consequence, the proportion of connected production units is simply a function

$$n^s(\kappa) \in [0, 1].$$

This function could be estimated from aggregate data on demand, capacity, and connections. For the moment, we only assume it is weakly increasing in κ, and we define the elasticity of the proportion of connected producers to the capacity index as

$$\nu^s(\kappa) = \frac{\kappa \frac{\partial n^s}{\partial \kappa}(\kappa)}{n^s(\kappa)}. \tag{4}$$

than $\ell(e)$ by concavity of v. Alternatively, one could possibly choose which consumers to disconnect: then one would disconnect those with the lowest individual value of lost load, creating a social loss that lies below $\ell(e)$. The point here is that the definition of the VOLL depends on assumptions on the management of demand. The notion of VOLL is thus more complex than it may seem at first glance. As discussed in Joskow and Tirole (2007), it also depends for example on whether consumers may undertake precautionary measures. Our definition also assumes that all consumers optimize their consumption and face the same unit price.

[14]For instance, in February 2021, many of the natural gas processing facilities in Texas were curtailed, resulting in an almost 50 percent decline in natural gas production, and in a significant amount of natural-gas fired generating capacity being made unavailable. See Wolak (2022) for a discussion.

	$m^s(\kappa) = 1 \leq n^s(\kappa)\kappa$	$m^s(\kappa) = n^s(\kappa)\kappa < 1$
$n^s(\kappa) = 1$	**no disconnection** $K > D$	**rolling blackout** $K = mD$
$n^s(\kappa) < 1$	**supply disconnection** $nK > D$	**systemic blackout** $nK = mD$

Table 1: m is the proportion of fully served consumers, n is the proportion of connected producers.

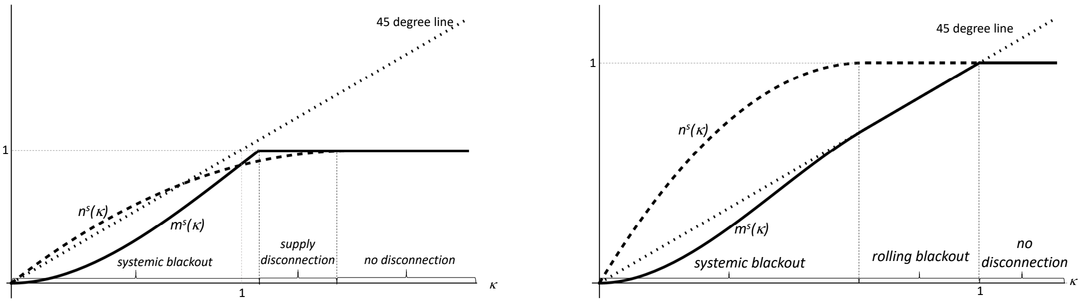
For the demand side, we assume that the grid is managed so as to serve as many consumers as possible. As discussed in the Introduction, this assumption seems reasonable, and in line with actual practice.[15] Since the connected capacity is nK and each consumer demands D, we obtain that the proportion m of connected consumers is exactly

$$m^s(\kappa) \equiv \min(n^s(\kappa)\kappa, 1). \tag{5}$$

Thus, the function n allows us to encompass a variety of possible regimes, that we describe in Table 2. When $m = 1$ (left column), supply is sufficient to satisfy demand, whereas in the right column ($m < 1$), the lack of capacity leads to blackouts. In the top-left box ("no disconnection"), all producers and consumers are connected to the grid, and connected capacity is sufficient to ensure that all demands are satisfied. From this regime, an incident on the grid (captured by the state of nature s), or a surge in demand, or a reduction in capacity (both impacting the capacity index κ), all may impact the proportion n of connected producers, and eventually the proportion m of served consumers. If the capacity index remains high enough, some producers may be disconnected while demand is still fully served (bottom left, "supply disconnection"). Alternatively, if n remains high while κ drops, some consumers have to be be disconnected even though producers remain connected ("rolling blackout"). The choice between these two paths thus depends on whether n or κ bear the brunt of the shock, as illustrated in Figures 1(a) and 1(b). Finally, more major shocks may impair the functioning of entire parts of the grid, and to

[15]In fact, in some cases it could be welfare-improving to disconnect more consumers than is needed, so as to make producers compete for the remaining demand. We leave this possibility for future research.

a "systemic blackout" in which both supply and demand is disconnected.



(a) Producers disconnected first.

(b) Consumers disconnected first.

Figure 1: Two possible paths as κ is reduced, for a given state s.

Note finally that thanks to the uniform disconnection assumption, each producer can be seen as endowed with the aggregate cost function C, and each of the n connected producers has to satisfy the demand of m/n consumers. Therefore, the realized total cost equals

$$nC^s\left(\frac{m}{n}D^s(\overline{p}), X\right).$$

3 Optima

In this Section, we characterize the allocations that maximize total surplus (some details of the derivations are gathered in the Proof Appendix). In our model, an allocation specifies the ex-ante investments, and the productions and consumptions in each state. Individual consumptions are constrained by rigid demand management: there must exist a retail price $\overline{p}$ such that in each state s consumption is $D^s(\overline{p})$, or zero if the consumer is disconnected. Let the variable $\kappa(s)$ be the value of κ in state s:

$$\kappa(s)D^s(\overline{p}) = \sum_k x_k K_k^s = K^s(X). \tag{6}$$

Remember that Definition (5) gives the proportion of consumers that are fully served, $m^s(\kappa) \equiv \min(n^s(\kappa)\kappa, 1)$. The consumers' gross surplus in state s is

$$m^s(\kappa(s))v^s(D^s(\overline{p})) + (1 - m^s(\kappa(s)))v^s(0),$$

from which we have to subtract the production cost

$$n^s(\kappa(s))C^s\left(\frac{m^s(\kappa(s))}{n^s(\kappa(s))}D^s(\overline{p}),X\right)$$

and the investment cost

$$\sum_k I_k(x_k).$$

Optimal allocations maximize the expectation over s of this aggregate surplus with respect to the capacity indexes $\kappa(s)$ in each state s, the retail price $\overline{p}$, and the investments (x_k), under Constraints (6). The multiplier $\beta(s)$ associated to each of these constraints thus measures the social value of one additional unit of capacity in state s, and we expect it to be positive.

Let us first provide necessary conditions associated to the choice of $\kappa(s)$. Note that for the sake of clarity, we often omit arguments in the formulas below. In states where all consumers are served ($m = 1$), the objective reduces to

$$v(D) - nC\left(\frac{D}{n},X\right),$$

to be maximized under Constraint (6). Using Definition (4) of the elasticity of n in , we obtain

$$\text{If } m = 1, \qquad \beta(s) = \frac{\nu}{\kappa}\Big[C_Q(\frac{D}{n}) - AC(\frac{D}{n})\Big]. \tag{7}$$

Here, the only value of an additional unit of capacity is to increase the capacity index and finally the number of connected producers, with the elasticity ν. This in turn reduces the production cost $nC(\frac{D}{n})$, to the extent that marginal cost is above the average cost. This effect is zero if all producers are connected (the no-disconnection regime in Table 2), or if producers are homogenous, and it is strictly positive otherwise (supply disconnection in Table 2).

The picture is more complex for states in which a blackout occurs ($m < 1$). Then the objective reduces to

$$n\kappa(v(D) - v(0)) - nC(\kappa D, X),$$

to be maximized under Constraint (6). After some simplifications, we obtain

$$\text{If } m < 1, \qquad \beta(s) = n\Big[\ell(D) - C_Q(K) + \nu(\ell(D) - AC(K))\Big]. \tag{8}$$

Now, the additional unit of capacity directly increases the proportion m of connected consumers, with a gain equal to the VOLL, at the price of a marginal increase in the cost (regime rolling blackout in Table 2). It also has an indirect effect, when the proportion of connected producers n is less than one (systemic blackout in Table 2): then a higher capacity index increases the number of connected producers n, and this allows to produce more and to serve more consumers, now at the average cost since the additional connected producers are drawn at random in the population of producers.

Provided that the VOLL ℓ is high enough compared to costs, we indeed obtain that the social value of additional capacity is positive: it allows to directly produce more, and it also indirectly makes the grid more stable, so as to connect more producers and more consumers. The second effect corresponds to a contribution to a public good, i.e. the stability of the grid. It is scaled by the value of the elasticity ν, which measures the endogeneity of blackouts.

Regarding investments, the first-order condition with respect to the number x_k of units of type k can be simplified using Property (1) to become (assuming it is optimal to use at least one such unit):

$$I_k'(x_k) = E\Big[n\pi_k(C_Q)\Big] + E[\beta K_k]. \tag{9}$$

The two first terms measure the private incentives to invest, assuming electricity is priced at the global marginal cost (though we shall soon explain why this is unlikely to be the case in equilibrium). The last term is the expected social value of this additional unit. We finally turn to the first-order condition with respect to the retail price $\overline{p}$, in which we use Definition (2) of the elasticity of demand in to get:

$$E\Big[\varepsilon m D(\overline{p} - C_Q)\Big] = E[\varepsilon \beta K]. \tag{10}$$

This means that the optimal retail price should equal an average of marginal costs, computed with a complex distribution that takes into account both total consumption mD and the elasticity of demand ε, to which must be added a term which measures the beneficial impact of a reduction in demand on the value of the capacity index.

4 Competitive Equilibria (with price cap and rationing)

When applied to wholesale electricity markets, the standard notion of competitive equilibria faces a difficulty: because retail pricing is rigid, no price can balance demand and supply when available capacity is less than realized demand, so that the idea of price-taking agents becomes awkward. The notion of competitive equilibria we shall use thus has to be augmented with regulatory interventions.

The first regulatory reaction to an imbalance on the market is to disconnect some consumers. But when the minimum number of consumers is disconnected, as we have assumed, every producer remains indispensable to the satisfaction of the remaining consumers. With non-infinitesimal producers, being indispensable means being able to increase prices without bounds. When observing such a surge in prices, the second regulatory reaction consists in imposing a price cap P, set ex-ante at a reasonable level to settle transactions between producers and retailers.

In our model, we introduce these two regulatory interventions as follows. In states where $m = 1$, there is excess capacity, producers act as price-takers, and electricity is priced at marginal cost, so that the wholesale price of electricity is

$$p(s) = C_Q^s\left(\frac{D}{n}, X\right). \tag{11}$$

Otherwise, when $m < 1$ every producer becomes indispensable and can extract an arbitrary high price for their production. In such a case, we require that the electricity price be capped at P, assumed high enough compared to costs so as to ensure that all available capacity is put to use:

$$p(s) = P > C_Q^s(K^s(X), X). \tag{12}$$

This modeling choice can be seen as a simple manner to model imperfect competition. Regrettably, it introduces a discontinuity in the wholesale price between states of the world in which $m < 1$ (blackouts) and states in which $m = 1$ (no blackouts). Notice however that this discontinuity would still appear if one were to model more precisely imperfect competition between producers. In such an hypothetical model, there would appear a threshold for m such that, when m is above this threshold, then pricing would obey the standard rules of, say, Cournot pricing, as for example in Elliott (2024); while if m is below the threshold, strategic producers would withdraw capacities so as to trigger the imposition of the price cap P. Whether such opportunistic behaviors may be discouraged by the fear of public reactions is difficult to assess; McRae and Wolak (2019) claim to have identified such behaviors for the Columbian electricity market. The point here is that this discontinuity is a natural consequence of the possibility of blackouts, and of strategic behavior on the supply side.

As discussed in the Introduction, a natural candidate for the cap is the VOLL, but various difficulties (not explicitly considered in our model) typically lead regulators to set the price cap below the VOLL. This is first to reduce the risk that retailers bear by buying at the risky wholesale price to sell at the fixed retail price. This is also to reduce the market power enjoyed by dominant producers (see Wolak (2013), and Fabra (2018), among many others). We shall thus take P as exogenously given, plausibly below the VOLL, but definitely above production costs so as to ensure that all available capacity is indeed used in case of blackouts, as stated in Property (12).

Still, setting the price cap below the VOLL creates two additional distortions, requiring two additional instruments. The well-known "missing money" problem underlines that with a low cap, producers are not rewarded sufficiently for their investments. One should thus introduce additional incentives for capacity creation. We therefore introduce a subsidy σ_k for each unit of type k. In addition, a low cap also reduces the price retailers have to pay for getting electricity, thereby reducing the competitive retail price. One has to correct this distortion by creating a tax τ on retail electricity.

This being clarified, let us now turn to the derivation of competitive equilibria, for given values of the price cap P, the retail tax τ, and the subsidies to capacity (σ_k). As before, one has to find an allocation in each state of nature, together with endogenous equilibrium prices which are the rigid retail price $\overline{p}$ and the flexible wholesale prices $p(s)$. Constraint (6) defining $\kappa(s)$ still holds, by definition of the variables involved. Wholesale prices $p(s)$ are determined in each state, according to Properties (11) and (12).

Recall that retailers buy electricity at the wholesale price, to resell it to consumers at the retail price, including the retail tax. Being competitive intermediaries with constant returns to scale, they take wholesale prices as given, and their expected profits are zero.

The retail contract specifies that consumers can buy whatever quantity they wish, in every state of nature, at the fixed price $\overline{p}$. This contract itself can be sold, for a fixed fee A. Competition should then select the contract $(A, \overline{p})$ that maximizes the consumer's payoff

$$E[mv^s(D^s(\overline{p})) - \overline{p}D^s(\overline{p})) + (1-m)v^s(0)] - A,$$

under the constraint that profits be nonnegative:

$$A + (\overline{p} - \tau)E[mD^s(\overline{p})] \geq E[p(s)mD^s(\overline{p})]. \tag{13}$$

By making this constraint bind, one obtains the classical result that the use of an access charge promotes efficiency, since the equilibrium contract gives zero-profit to retailers and maximizes over $\overline{p}$ the surplus

$$E\Big[m[v^s(D^s(\overline{p})) - (p(s) + \tau)D^s(\overline{p})] + (1-m)v(0,s)\Big].$$

This result appears also in (Joskow and Tirole, 2006, 2007). These articles also underline that retail contracts raise difficulties of their own when consumers are heterogenous, since different consumers with different load profiles entail different expected costs for the retailer.[16] For the sake of simplicity, we abstract from these difficulties, and we only consider one retail contract, for which $\overline{p}$ maximizes the expected aggregate surplus, with the following first-order condition:

$$(\overline{p} - \tau)E[\varepsilon mD] = E[p(s)\varepsilon mD]. \tag{14}$$

Finally, every investment should balance the marginal cost of an additional unit with the additional expected revenues, so that, once more assuming an interior solution, and taking the capacity subsidy into account:

$$\text{for all } k \qquad I'_k(x_k) = E[n\pi^s_k(p(s))] + \sigma_k. \tag{15}$$

[16] This implies in particular that the use of such contracts creates an adverse selection problem for the retailers, as the profits made on each consumer depend on its individual load profile–*i.e.*, on whether they consume electricity at peak hours or not. We ignore this problem in this article, and we refer to Joskow and Tirole (2006) for a related study, and to Leslie, Pourkhanali and Roger (2021) for an estimate of cross-subsidies between consumers that sign the same contract.

5 Decentralization

There remains to determine the values of regulatory instruments that decentralize a given optimal allocation, for a given price cap P. To state the result more clearly, we first define in each state the following measure of missing incentives:

$$\Delta(s) = \begin{cases} n(\ell(D) - P) + \nu n\,(\ell(D) - AC(K)) & \text{for } m < 1 \quad (16a) \\ \dfrac{\nu}{\kappa}\left(C_Q(\dfrac{D}{n}) - AC(\dfrac{D}{n})\right) & \text{for } m = 1. \quad (16b) \end{cases}$$

In case of blackouts ($m < 1$), the first term is associated to the "missing money" problem: agents trade at the price cap instead of facing the social value of electricity, which equals the VOLL ℓ. The second term which is weighted by ν is associated to the "public good" problem: additional capacity allows to limit rationing. Without blackouts ($m = 1$) when demand is fully served, only the second term remains. The additional capacity reduces the total production cost. Observe that

$$\Delta = \beta - 1_{m<1} n(P - C_Q(K)), \tag{17}$$

so that the missing incentives equal the social value of capacity β defined in (7) and (8), minus a term in case of blackout, due to the fact that the price cap corrects part of the gap between the social value ℓ and the marginal cost. We are now ready to state our main result:

Proposition 1 *Let the price cap P be given. To decentralize an optimal allocation, one needs to set the retail tax τ and the capacity subsidies σ_k as follows:*

$$\tau E[\varepsilon m D] = E[\varepsilon \Delta K] \tag{18}$$

$$\textit{For all } k, \qquad \sigma_k = E[\Delta K_k]. \tag{19}$$

The proof of this result simply relies on comparing the Equations (7)-(8)-(9)-(10) characterizing an optimum, to the Equations (11)-(12)-(14)-(15) characterizing an equilibrium. The resulting values for the tax and the subsidies thus constitute only necessary conditions. Still, they yield interesting insights.

Instruments are sufficient. The first result is that even though the price cap is set at an arbitrary level, the other instruments are flexible enough to decentralize an optimal allocation. In fact, both the missing money problem and the public good problem are tackled by the same instruments: this is why Δ is a sum of terms, each originating from one of these problems. This is a nice normative result, which complements the positive analysis in Joskow and Tirole (2007).

A retail tax is needed. Notice moreover that both the retail tax and the capacity subsidies are decreasing with the price cap; for the retail tax, observe that the slope lies between -1 and 0.[17] In particular, one may for example find the value of P for which the retail tax is zero: we obtain

$$PE[\varepsilon nK] = E[\varepsilon nK(\ell(D) + \nu(\ell(D) - AC(K))],$$

which means that one should set the price cap above the expected Value of Lost Load, a very high level indeed. The model thus supports the creation of a positive retail tax, simply because by reducing demand it alleviates both the missing money problem and the public good problem.

Subsidies to capacities are not uniform. Despite their apparent complexity, these formulas may be used in various directions. For example, one may wonder in which case one can optimally use a uniform subsidy per unit of capacity. To answer this question, let us focus on the case when **capacities are affected by the same multiplicative shock** h:

$$K_k^s = h(s)\overline{K}_k. \tag{20}$$

Then Equation (19) characterizes an optimal subsidy per unit of capacity which is indeed independent from the unit type:

$$\frac{\sigma_k}{\overline{K}_k} = E[nh(s)\Delta].$$

One may also allow for independent shocks on capacity, for example due to maintenance operations, and rely on the law of large numbers to get the same result. But these

[17]Simply because $E[\varepsilon mD]\frac{\partial \tau}{\partial P} = -E[1_{m<1}\varepsilon nK] = -E[1_{m<1}\varepsilon mD] \geq -E[\varepsilon mD]$.

are very restrictive cases, and we can conclude that in the presence of intermittency, optimal subsidies to capacity cannot be uniform.

In the general case, optimal subsidies simply give a value to capacity. This capacity varies according to breakdowns or maintenance operations for nuclear plants and thermal production from fossil fuels, but it is much more volatile for renewable sources that depend on weather. The computation of the expectation in (19) is thus much more complex, as it should take into account the correlation between the missing incentives and the effective capacity. Our formulas indicate how such a computation can be performed, using statistical information from past chronicles.

When the regulator's budget is balanced. In addition, we are able to obtain striking results in a particular case often used in the theoretical literature, as well as in empirical studies of the electricity sector[18]: the case when **demand is impacted by a multiplicative shock** b, so that:

$$D(p, s) = b(s)d(p), \tag{21}$$

where $b > 0$ and the function d is decreasing. Under this assumption, a number of straightforward results follow (see Section C Appendix for details of the derivations):

i) The competitive access fee A is zero;

ii) The elasticity of demand $\varepsilon^s(\overline{p})$ and the Value of Lost Load $\ell^s(D^s(\overline{p}))$ are both independent from s;

iii) The revenues from the retail tax exactly balance the cost of subsidies:

$$\tau EmD = \sum_k x_k \sigma_k.$$

The assumption of a multiplicative shock does not seem very restrictive in practice, and it allows to give a clear meaning to the VOLL. Remarkably, it also yields the result that the regulatory budget is exactly balanced. This remarkable identity is new in the literature, and may be useful as a guide for regulatory policies.

[18]Recent papers such as **?** typically assume a multiplicative shock, and an iso-elastic demand function. This last assumption is in fact not needed to establish our results.

The price cap is not enough. Finally, some works mention the idea that a price cap is a sufficient instrument, provided it is set at the right level, i.e., the VOLL.[19] Our model highlights how restrictive the conditions are for this result to hold: one needs that capacities are affected by the same multiplicative shock (no intermittency, as in (20), and that demand is affected by a multiplicative shock (as in (21), and that the proportion of connected producers does not depend on the capacity index ($\nu = 0$), meaning that this proportion is exogenous. Then one can set the retail tax and the capacity subsidies to zero, provided the price cap is set at the VOLL. In all other cases, the price cap has to be complemented by other regulatory instruments: the energy-only paradigm seems to rely on a weak basis.

6 Conclusion

We have set up a simple but general model of the electricity sector. We have focused on two issues: the inability to balance the market creates the need for rationing and for introducing a price cap, which in turn creates a missing money problem; and the public good problem, associated to the reliability of the grid, which depends on the excess capacity relative to demand. We have shown that these two issues can be tackled efficiently thanks to a retail tax, and to differentiated subsidies to production. Under the assumption that demand is only affected by a multiplicative shock, we have established that it is optimal to balance the budget of the regulatory agency.

We thus provide a simple and effective regulation. Whether it is robust to market power remains to be studied. On the other hand, it is easily understood that all these regulatory efforts stem from the inability to regulate demand in real-time. Any progress in this direction would not only reduce the probability of rationing, and the inefficiencies associated to the price cap; it would also reduce market power, by making demand more elastic. Real-time pricing is thus likely to be the relevant frontier for efficiency gains, and for alleviating the weight of regulation.

[19] This is sometimes called the energy-only paradigm. For recent references, see Keppler (2017), Fabra (2018), and Keppler, Quemin and Saguan (2022).

A Details in the derivations in Section 2

Total expected surplus is

$$E\left[m^s(\kappa(s))v^s(D^s(\overline{p})) + (1-m^s(\kappa(s)))v^s(0) - n^s(\kappa(s))C^s\left(\frac{m^s(\kappa(s))}{n^s(\kappa(s))}D^s(\overline{p}), X\right)\right] - \sum_k I_k(x_k)$$

that we maximize wrt $(\kappa, \overline{p}, X)$, under the constraints (with multiplier $\beta(s)$ in each state)

$$\kappa(s)D^s(\overline{p}) = \sum_k x_k K_k^s = K^s(X).$$

Let us first provide necessary conditions associated to the choice of $\kappa(s)$. Note that for the sake of clarity, we often omit arguments in the formulas below. In states where all consumers are served ($m = 1$), the Lagrangian reduces to

$$v(D) - n(\kappa)C\left(\frac{D}{n(\kappa)}, X\right) - \beta\kappa D.$$

Use the identity $\frac{\partial n}{\partial \kappa} = \nu n/\kappa$ from (4), and divide by D to obtain Equation (7):

$$\text{If } m = 1, \qquad \beta(s) = \frac{\nu}{\kappa}\Big[C_Q(\frac{D}{n}) - AC(\frac{D}{n})\Big].$$

In states in which a blackout occurs ($m < 1$), the Lagrangian reduces to

$$n(\kappa)\kappa(v(D) - v(0)) - n(\kappa)C(\kappa D, X) - \beta\kappa D.$$

Use once more the identity $\frac{\partial n}{\partial \kappa} = \nu n/\kappa$, and divide by D to obtain Equation (8):

$$\text{If } m < 1, \qquad \beta(s) = n\Big[\ell(D) - C_Q(K) + \nu(\ell(D) - AC(K))\Big].$$

Regarding investments, the first-order condition with respect to the number x_k of units of type k can be simplified using Property (1) to derive Equation (9):

$$I_k'(x_k) = E\Big[n\pi_k(C_Q)\Big] + E[\beta K_k].$$

We finally turn to the first-order condition with respect to the retail price $\overline{p}$, in which we use the identities $D_p = -\varepsilon D/\overline{p}$ and $\kappa D = K$ to get Equation (10):

$$E\Big[\varepsilon m D(\overline{p} - C_Q)\Big] = E[\varepsilon\beta K].$$

This concludes this part. □

B Proof of Proposition 1

Equation (10) characterizing the optimal retail price reads

$$E\Big[\varepsilon m D(\overline{p} - C_Q)\Big] = E[\varepsilon \beta K].$$

The left-hand side can be decomposed into

$$E\Big[\varepsilon m D(\overline{p} - p)\Big] + E\Big[\varepsilon m D(p - C_Q)\Big].$$

From Properties (11) and (12) for the wholesale price p, the second term is zero when $m = 1$; and when $m < 1$, then $mD = nK$ and $p = P$, so that this second term reduces to

$$E\Big[1_{m<1} \varepsilon n K(P - C_Q(K))\Big],$$

so that we have:

$$\overline{p} E\Big[\varepsilon m D\Big] + E\Big[1_{m<1} \varepsilon n K(P - C_Q(K))\Big] = E[\varepsilon \beta K] + E\Big[p \varepsilon m D\Big].$$

Comparing to Equation (14), we obtain the value of the optimal retail tax:

$$\tau E\Big[\varepsilon m D\Big] = E\Big[\varepsilon \beta K\Big] - E\Big[1_{m<1} \varepsilon n K(P - C_Q(K))\Big].$$

The result in the Proposition follows, using Definition (17).

We proceed similarly for the optimal subsidies. For each type k of production unit, Equation (9) characterizing the optimal investments reads:

$$I_k'(x_k) = E\Big[n \pi_k(C_Q)\Big] + E[\beta K_k].$$

Let us first analyze profits. When $m = 1$, from Property (11), we have $p(s) = C_Q$, so that

$$\pi_k(C_Q) = \pi_k(p(s)).$$

When $m < 1$, we have $p(s) = P$, and the unit produces at full capacity both when the price is P and when the price is $C_Q(K)$. Therefore,

$$\pi_k(C_Q(K)) = C_Q(K) K_k - c_k(K_k) = (C_Q(K) - P) K_k + \pi_k(p(s)).$$

Replacing, we obtain

$$I'_k(x_k) = E\Big[n\pi_k(p(s))\Big] + E[\beta K_k] - E\Big[1_{m<1} n(P - C_Q(K))K_k\Big].$$

Compare to the Equation (15) characterizing the equilibrium investments to get

$$\sigma_k = E[\beta K_k] - E\Big[1_{m<1} n(P - C_Q(K))K_k\Big] = E[\Delta K_k].$$

This concludes the proof. □

C A study of the case when demand is impacted by a multiplicative shock

Recall the definition of this case in Definition (21):

$$D(p, s) = b(s)d(p),$$

where $b > 0$ and the function d is decreasing. This demand function derives from the following surplus function

$$v(e, s) = a(s) + b(s)G\left(\frac{e}{b(s)}\right),$$

where G is a primitive of the inverse function d^{-1} of d. Now, it is easily seen that the elasticity of demand

$$\varepsilon^s(p) = \frac{-pd'(p)}{d(p)}$$

and the VOLL measured at the consumption point

$$\ell^s(D^s(p)) = \frac{G(d(p)) - G(0)}{d(p)}$$

are both independent of s, as announced. Moreover, ε simplifies in Equation (14), so that Constraint (13) implies that the competitive access charge is exactly zero, as announced. Finally, ε also simplifies in Equation (18), which reduces to

$$\tau E[mD] = E[\Delta K],$$

and the right-hand side is exactly $\sum_k \sigma_k x_k$, from Equation (19). This proves that the regulatory budget is balanced, as announced.

References

Ambec, Stefan, and Claude Crampes. 2019. “Decarbonizing electricity generation with intermittent sources of energy.” *Journal of the Association of Environmental and Resource Economists*, 6(6): 1105–1134.

Borenstein, S, and S Holland. 2005. “On the efficiency of competitive electricity markets with time-invariant retail prices. RAND Journal of Economics 36 (3): 469-493.”

Borenstein, Severin, James Bushnell, and Erin Mansur. 2023. “The economics of electricity reliability.” *Journal of Economic Perspectives*, 37(4): 181–206.

Butters, R Andrew, Jackson Dorsey, and Gautam Gowrisankaran. 2024. “Soaking up the sun: Battery investment, renewable energy, and market equilibrium.” National Bureau of Economic Research.

Elliott, Jonathan T. 2024. “Investment, Emissions, and Reliability in Electricity Markets.”

Fabra, Natalia. 2018. “A primer on capacity mechanisms.” *Energy Economics*, 75: 323–335.

Ferrasse, Jean-Henry, Nandeeta Neerunjun, and Hubert Stahn. 2022. “Intermittency and electricity retailing: An incomplete market approach.” *Mathematical Social Sciences*, 120: 24–36.

Gerlagh, Reyer, Matti Liski, and Iivo Vehviläinen. 2022. “Rational Rationing: A Price-Control Mechanism for a Persistent Supply Shock.”

Gorman, Will. 2022. “The quest to quantify the value of lost load: A critical review of the economics of power outages.” *The Electricity Journal*, 35(8): 107187.

Gowrisankaran, Gautam, Ashley Langer, and Wendan Zhang. 2024. “Policy Uncertainty in the Market for Coal Electricity: The Case of Air Toxics Standards.”

Gowrisankaran, Gautam, Stanley S Reynolds, and Mario Samano. 2016. "Intermittency and the value of renewable energy." *Journal of Political Economy*, 124(4): 1187–1234.

Joskow, Paul, and Jean Tirole. 2006. "Retail electricity competition." *The RAND Journal of Economics*, 37(4): 799–815.

Joskow, Paul, and Jean Tirole. 2007. "Reliability and competitive electricity markets." *The RAND Journal of Economics*, 38(1): 60–84.

Joskow, Paul L. 2022. "From hierarchies to markets and partially back again in electricity: responding to decarbonization and security of supply goals." *Journal of Institutional Economics*, 18(2): 313–329.

Keppler, Jan Horst. 2017. "Rationales for capacity remuneration mechanisms: Security of supply externalities and asymmetric investment incentives." *Energy Policy*, 105: 562–570.

Keppler, Jan Horst, Simon Quemin, and Marcelo Saguan. 2022. "Why the sustainable provision of low-carbon electricity needs hybrid markets." *Energy Policy*, 171: 113273.

Léautier, Thomas-Olivier. 2016. "The visible hand: ensuring optimal investment in electric power generation." *The Energy Journal*, 37(2): 89–109.

Leslie, Gordon, Armin Pourkhanali, and Guillaume Roger. 2021. "Can real-time pricing be progressive? Identifying cross-subsidies under fixed-rate electricity tariffs." *Identifying Cross-Subsidies under Fixed-Rate Electricity Tariffs (January 27, 2021).*

Llobet, Gerard, and Jorge Padilla. 2018. "Conventional Power Plants in Liberalized Electricity Marketswith Renewable Entry." *The Energy Journal*, 39(3): 69–92.

McRae, Shaun D, and Frank A Wolak. 2019. "Market power and incentive-based capacity payment mechanisms." *Unpublished manuscript, Stanford University.*

Wolak, Frank A. 2013. "Economic and political constraints on the demand-side of electricity industry re-structuring processes." *Review of Economics and Institutions*, 4(1): 42.

Wolak, Frank A. 2022. "Long-term resource adequacy in wholesale electricity markets with significant intermittent renewables." *Environmental and Energy Policy and the Economy*, 3(1): 155–220.