Long-term issues with the energy-only market design in the context of electricity decarbonization: Insights from a system dynamics simulation model

Alexis Lebeau, Marie Petitet, Simon Quemin, Marcelo Saguan

Long-Term Issues with the Energy-Only Market Design in the Context of Deep Decarbonization

Alexis Lebeau*[,1,2], Marie Petitet[1,3], Simon Quemin[1,4,5], and Marcelo Saguan[1]

[1] Électricité de France R&D, EDF Lab Paris-Sacaly, France

[2] Industrial Engineering Research Department, CentraleSupélec, France

[3] Chaire European Electricity Markets, University Paris-Dauphine, France

[4] Potsdam Institute for Climate Impact Research, Germany

[5] Grantham Research Institute, London School of Economics, United Kingdom

This version: July 2023

Abstract

There has been fierce controversy in the literature over the long-run efficiency of the energy-only market (EOM) design ever since its inception. In this paper, we provide novel insights to illuminate this historical controversy, and we revisit it with a focus on contemporary issues and profound changes brought about by the energy transition. Specifically, we develop an analytical and modeling framework to quantitatively investigate how EOM outcomes hinge on the underlying behavioral, informational and structural assumptions. We apply our framework to a case study calibrated on Californian fundamentals that captures the key features of energy systems under deep decarbonization. We characterize how EOM outcomes can substantially deviate from the long-run optimum as soon as one assumption is relaxed compared to theoretical requirements. This leads to pathways with higher electricity prices, lower security of supply and higher emission levels that imperil decarbonization. In particular, we highlight how market price signals alone are prone to a dynamic entry-exit coordination problem between investment in low-carbon assets and the phaseout of fossil-fired assets. This calls for a market design reform to complement price signals that accounts for realistic assumptions.

Keywords— Electricity Market Design, Energy-Only Market, Investments and Retirements, Decarbonization, System Dynamics, Generation Expansion Planning.

JEL codes— C63, D47, D81, L94, P18.

*Corresponding author: `alexis.lebeau@edf.fr`

1 Introduction

The energy transition poses mounting challenges to energy systems across the world with various interrelated facets including decarbonization, renewables integration, energy efficiency improvement and electrification. In practice, these ambitious aspirations translate into different targets at different horizons. For instance, Senate Bill 100 in California mandates targets of 60% and 100% zero-carbon electricity retail sales to end-use customers by 2030 and 2045, respectively (California State Senate, 2018). The European Union Green Deal and fit-for-55 policy package are another case in point, with the 2030 target of cutting greenhouse gas emissions by 55% below 1990 levels and the objective of a net-zero economy by mid-century (European Commission, 2019, 2022).

In the electricity sector, generation expansion planning (GEP) models are often used to explore cost-efficient pathways to achieve decarbonization objectives. There is a rich literature discussing the underlying optimization techniques and technological assumptions (e.g., integration of renewable energy sources, representation of short-term operations and seasonal storage) in such analyses (e.g., Alimou et al., 2020; Abdin et al., 2022). There are also numerous applications that provide key insights on decarbonization pathways in given jurisdictions – e.g., CPUC (2019) for California or RTE (2021) for France – or on the role of specific technologies – e.g., hydrogen in Schulthoff et al. (2021). This class of models offers a normative framework that is widely used by academics, regulators, agencies, consultancies, investors and market participants alike.

A crucial assumption of these models is that they take the perspective of a benevolent social planner. While this has the advantage of yielding normative results that hold independently of the institutional or market framework, this turns into a blind spot when the latter is to be studied and scrutinized. In particular, there is increasing concern that the energy-only market (EOM) design,[1] which has been held up as the target design model in many jurisdictions since the beginning of the deregulation era, may fall short of supporting necessary investments to deliver decarbonization, reliability and affordability objectives in a cost-efficient and timely manner (e.g., Roques and Finon, 2017; Newbery, 2018; Blaquez et al., 2020; Joskow, 2022; Keppler et al., 2022).

The EOM paradigm rests on the principle that socially optimal long-term entry and exit decisions can be decentralized in competitive markets. Theoretical foundations can be traced back to Arrow and Debreu (1954) in a general framework and to Caramanis (1982) for a seminal application to the electricity industry.[2] They are based on the equivalence between optimality conditions

[1] The traditional 'energy-only' terminology used in the literature can sometimes be misleading as it may include a complete sequence of short-term markets (e.g., for adjustment and ancillary services) as well as derivatives (e.g., futures contracts) markets. In this paper, we use this terminology to refer to fully deregulated market designs based on short-term wholesale markets and associated derivatives markets.

[2] It is worth recalling that Boiteux (1949, 1960) established seminal results on marginal electricity pricing for a regulated utility under the assumption of an optimal investment policy: "*Provided there is an optimal investment policy, short-term pricing is also long-term pricing, and there is no longer any contradiction between the two.*" Similarly, Schweppe et al. (1988) also considered an energy marketplace in a regulated environment. Acknowledging the investment decentralization result of their co-author (Caramanis), Schweppe et al. wrote "*The spot price based energy marketplace is designed to operate in a regulated environment [...]. This chapter only presents a set of basic ideas [...]. Since the advantages and disadvantages have not been quantified, we are not advocating deregulation (i.e., we do not know whether there is 'a lady or a tiger' behind the door).*"

of private agents in a perfectly competitive market and those of a benevolent social planner (also assuming private and social discount rates are equal). From an investor perspective, this notably translates into the equality between inframarginal rents and fixed costs at the optimum. Although commonly presented for a representative year with annualized costs under certainty (e.g., Joskow and Léautier, 2021), this result can be extended to a multi-year, stochastic framework by equating the expected discounted sum of inframarginal rents with total fixed costs (e.g., Poncelet et al., 2019). Yet this equivalence holds under "*several strong simplifying hypotheses*" (Rodilla and Batlle, 2012) or "*demanding conditions*" (Newbery, 2018), including perfect information, full rationality of agents, and complete markets for risk trading (see Section 4.1.1 for detail).

Whether or not the tenability of these assumptions would compromise the practicability of the EOM paradigm has spurred heated debates in the community since its very inception.[3] At the dawn of the deregulation era, Littlechild (1988) observed the disagreement, noting that "*mathematical models designed to prove that spot pricing is socially optimal are unpersuasive*" – referring to the aforementioned "*MIT models*" of Caramanis (1982) and Schweppe et al. (1988). More skeptical of the purported long-run efficiency of spot markets, others like Westfield (1988) quite remarkably foreshadowed detrimental impacts on the cost of capital that are front and center in current debates (e.g., Peluchon, 2021; Gohdes et al., 2022; Neuhoff et al., 2022; Newbery, 2023).[4] Even though the theoretical controversy never settled, political impetus was a critical driver for the liberalization of the industry (e.g., Joskow and Schmalensee, 1988; Léautier, 2019). The implications of deregulation for long-run efficiency also took a long time to materialize in practice, notably because wholesale markets were implemented in relatively mature power systems with a stable or contracting demand and little need for new investments. The main focus was on short-run efficiency, i.e. on harnessing market forces to ensure an efficient use of the existing fleet (e.g., Pollitt, 2021; Cicala, 2022).[5]

Over the last decade or so, this controversy has been reinvigorated in three phases with new variations. First, as documented in Bublitz et al. (2019), security of supply and 'missing money' in liberalized markets gradually became a focus of attention, with debates on the need for and design of capacity remuneration mechanisms. Second, as documented in Keppler et al. (2022), the energy transition shed a new light on the debates due to the required profound changes in energy system structures and generation mixes. In particular, the sheer scale and speed of necessary low-carbon investments, their specificities and capital-intensiveness, and various externalities (e.g., learning spillovers, social and industrial preferences) generate a market design and regulatory conundrum. While these problems are now rather well delineated at a conceptual level, there is scant literature

[3] This prolonged a similar controversy in the public utility pricing literature on the conditions for the equivalence between short– and long-run marginal cost pricing, see e.g. Andersson and Bohman (1985) and Section 4.1.

[4] More generally on long-run efficiency, Westfield (1988) warned that "*spot markets for electric power will not perform the miracles that perfect markets perform in the economic theory textbook. Many of the gains achievable through centralized coordination will be lost.*"

[5] Back then the context and policy objectives were markedly different from those that prevail today for the energy transition. The focus was on improving the efficiency of mature systems and on replacing old coal-fired plants with modern gas-fired plants as peaking units (financing needs were modest because of CCGT's low fixed costs relative to variable costs). See also Section 5 in Keppler et al. (2022) for a short historical perspective and discussion.

offering quantitative insights.[6] Last but not least, the ongoing energy crisis further exposed pre-existing design shortcomings and initiated a new wave of market reforms, also putting the question of affordability at the core of the debates (e.g., Fabra, 2023; Schittekatte and Batlle, 2023).

The objective of this paper is to provide novel quantitative elements to illuminate this historical controversy on electricity market design and to revisit it with a focus on contemporary issues and profound changes brought about by the energy transition. Specifically, we develop an analytical and modeling framework to quantitatively investigate how EOM outcomes are sensitive to theoretical assumptions and characterize how they deviate from the long-run optimum when these assumptions are relaxed. We apply our framework to a case study capturing the key features of energy systems under deep decarbonization, which differ from those when short-term markets were introduced and call for specific investigation. More precisely, we make three contributions to the literature.

As a first contribution, we provide an analytical framework to unpack the underlying assumptions that govern the efficiency of an EOM design in the long run. Specifically, we clearly delineate the behavioral, informational and structural assumptions that are conducive to an optimal energy mix in a pure EOM. In particular, we emphasize the dynamic nature of entry-exit decisions and the crucial roles of risks, hedging and anticipations. Compared to the existing literature that has looked into these assumptions at a conceptual level (e.g., Joskow, 2008; Rodilla and Batlle, 2012; Newbery, 2018) or with a focus on specific assumptions (e.g., Kraan et al., 2019; Fraunholz et al., 2023; Tao et al., 2023), our framework is synthetic and unified. This allows us to relax assumptions separately in order to isolate and compare the effects of doing so.

As a second contribution, we develop a modeling framework combining optimization and simulation models in a novel way in order to operationalize the above analytical framework. Specifically, our core simulation model uses System Dynamics (SD) as a modeling approach to consider a representative agent and capture the aggregate market impact of relaxing each assumption. Compared to the literature that has used this or similar modeling approaches to analyze agent behavior and market design issues (see Section 2.1 for a detailed review), our SD market simulation model has key distinguishing features. First, it has a linkage with a GEP model that can be used to define the anticipations of the representative agent. The outcomes of the GEP model are also used as a normative benchmark against which the performances of simulated EOM outcomes are evaluated. Second, it also has a linkage with an optimization-based merit-order dispatch model to represent short-term operations and determine future wholesale electricity prices that are the only source of remuneration for assets in an EOM. Third, it solves for both investment and retirement decisions simultaneously in conventional, renewable and storage technologies.

As a third contribution, we combine these two frameworks and provide quantitative illustrations in the context of the energy transition. Specifically, we build a case study on the basis of a stylized

[6] To our knowledge, there are two exceptions: Kraan et al. (2019), who find that an EOM does not give sufficient and stable investment signals to sustain a renewable, reliable and affordable power system; Zimmermann and Keles (2023), who find that carbon neutrality cannot be achieved only through market-based investments in France.

representation of Californian fundamentals over 2025–45. We select the Californian system because of public data availability and reliability, and because it presents the key characteristics of power systems under decarbonization in a relatively simple setting.[7] This facilitates the modeling and the interpretation of the results, whose identified trends are relevant for all power systems in transition. These characteristics are a commitment to eliminate emissions that requires massive investments in non-dispatchable renewable energy sources along with storage solutions, an existing fleet with a sizable share of fossil-fired plants that will partly be phased out before the end of their lifespans, and an increasing demand at wholesale level driven by electrification.

Overall, our quantitative results illustrate how the theoretical assumptions needed for an EOM to deliver the long-run optimum can hardly be met in practice. In particular, we highlight the high level of informational and computational complexity associated with rational decision-making and optimal anticipations of all relevant market fundamentals and future entries and exits. Our results also reveal that EOM outcomes can substantially deviate from the optimum when we introduce limited anticipation sophistication or risk aversion that affect expected revenue streams and asset profitability.[8] This leads to higher electricity prices, lower security of supply and higher emissions. A crucial new insight from our results is a coordination problem between investment in low-carbon assets and the phaseout of fossil assets that hinders the energy transition. For instance, under risk aversion, our results go beyond the standard under-investment result established in the literature. Specifically, risk aversion results in delays in both new investment and fossil phaseout that mutually reinforce one another, and leads to emissions in excess of decarbonization targets.

The remainder proceeds as follows. Section 2 presents the modeling approach in relation with the literature and describes the modeling framework, notably the SD model that simulates EOM outcomes. Section 3 presents our case study, describes the model calibration and characterizes the long-run optimum obtained with the GEP model. Section 4.1 presents our analytical framework, examines the assumptions under which the SD-EOM model replicates the GEP optimal outcomes, and discusses the impacts of unit indivisibility. Section 4.2 explores how simulated EOM outcomes deviate from the optimum when we relax these assumptions. Section 4.3 summarizes our results and offers implications for policy and market design. Section 5 concludes and outlines how our modeling framework can be extended to assess and compare alternative designs of long-term contracting mechanisms currently contemplated as part of undergoing market reforms.

[7] For instance, imports and exports are relatively small, which makes the isolated single-zone assumption acceptable, and only two technologies – solar PV and battery storage – are envisaged as key transition drivers, attracting the bulk of new capacity investments by mid century (around 95% in 2045 according to CPUC, 2019), which reduces the set of technologies one needs to represent. See also Section 3.1 for detail on open data sources.

[8] As Section 4.1.1 will make clear, risk aversion is not a problem per se. It only distorts entry-exit decisions when markets are incomplete and all relevant risks cannot be traded, which we implicitly assume is the case here.

2 Modeling framework

In this section, we first provide an overview of our modeling framework and place it within the related literature. We next describe its two constituent models. Models were developed and coded in Python, and the open-source codes are provided here: GitHub/ANTIGONE.

2.1 Overview

Classes of models in the literature. For decades, electricity economists and engineers have used a rich toolbox of complementary approaches to model and get to grips with long-term power system issues. The various modeling options are generally classified into three categories, namely optimization, equilibrium, and simulation models (e.g., Ventosa et al., 2005; Creti and Fontini, 2019) with distinct and complementary areas of relevance:

- Optimization models are the original and traditional approach to modeling the evolution of energy systems. The so-called generation expansion planning (GEP) models typically take the perspective of a central planner that seeks to determine the socially optimal capacity development plan (i.e., that which minimizes system-wide investment and operating costs) given a variety of constraints (e.g., a cap on carbon emissions), see Kagiannas et al. (2004) for a historical perspective. Over time, GEP models have notably been extended to stochastic frameworks and are still widely used today to analyze decarbonization pathways for energy systems, see Weber et al. (2021) for a recent review.
- Equilibrium models simultaneously solve individual profit maximization problems for different types of agents (e.g., producers with different technologies, intermediaries, consumers), finding equilibrium solutions where no agent is better off deviating unilaterally (e.g., Fan et al., 2012). These models typically feature uncertainty and risk aversion (e.g., Ehrenmann and Smeers, 2011; Abada et al., 2017; Mays et al., 2019; Mays and Jenkins, 2022).
- Simulation models can represent different decision-making rules (i.e., beyond profit maximization) and degrees of agent's sophistication and rationality. There are two broad types of simulation models: The first is agent-based modeling (ABM) that can feature heterogeneous agents. The second uses system dynamics (SD) and typically considers representative agents. dos Santos and Saraiva (2021) and Tao et al. (2021) (resp. Teufel et al. (2013) and Ahmad et al. (2016)) provide useful reviews of ABM (resp. SD models) applied to energy systems.

Optimization models abstract the market realities away, and as such, they provide insightful and normative results. In fact, linear/convex GEP model results can be interpreted as the outcomes from perfectly competitive markets with fully rational and informed agents. While this constitutes a useful benchmark, it does not capture the market environment in which agents trade and invest, the individual decision-making process of market participants (possibly with bounded rationality, information or foresight) or the sequentiality of discrete investment/divestment decisions over time (since all time steps are solved simultaneously) – see also Section 4.1.1.

Equilibrium models allow modelers to represent and assess the impacts of the market structure and heteregeneous agents and behaviors. This notably endogenizes key decisions and model variables (e.g., risk trading and the associated cost of capital). Yet these models rely on solvers whose results do not lend themselves to a straightforward interpretation of the mechanisms leading to the equilibrium, and they are typically solved in steady state, which does not unveil the dynamic nature of investment/divestment decisions. Additionally, by design these models cannot account for out-of-equilibrium situations, which are acknowledged to be common and deserve more attention (e.g., de Vries and Heijnen, 2008; Léautier, 2019).

Compared to equilibrium models, simulation models give more latitude in making explicit assumptions about agents' rationality, information and foresight levels, and in representing out-of-equilibrium situations. Arguably, this strength can also be a weakness, in that assumptions must be clearly spelt out and articulated with one another in order to arrive at sensible modeling results. Although SD models were first developed and applied to the electricity sector (e.g., Ford, 1983; Bunn and Dyner, 1996), both ABM and SD models are widely used today in the context of the energy transition – see inter alia Keles et al. (2016), Kraan et al. (2019), Fraunholz et al. (2021, 2023), Tao et al. (2021, 2023), Anwar et al. (2022) for ABM, and Petitet et al. (2016, 2017), Ousman Abani et al. (2018), Rios-Festner et al. (2019, 2020), Tang et al. (2021) and Pourramezan and Samadi (2023) for SD models.

Modeling approach in this paper. We develop a modeling framework consisting at its core of an SD model that we complement by a GEP model. Given that we aim to assess the impacts of relaxing the assumptions underpinning the optimality of long-term decision-making in a canonical energy-only market (see Section 4.1.1), the choice of an SD model proceeds in two steps:

(1) We first opt for a simulation model because we wish to explicitly represent and vary the assumptions about investor behavior (i.e., rationality, foresight, information, risk aversion) as well as quantify and describe the temporal dynamics of the energy transition.

(2) We next select an SD model to focus on a representative investor and isolate the impacts of relaxing said assumptions at an aggregate level – i.e., possibly capturing the resultant of different agents' decisions, but without formally accounting for heterogeneous agents and behaviors.[9]

As discussed above, great care will be taken to motivate and delineate our assumptions about investor behavior and describe how we implement them in the following. To name but one key issue at this stage, Tao et al. (2021) and Fraunholz et al. (2021) note how simulation model results can be sensitive to (long-term) price projection methods, and in particular to the way future capacity

[9] Given that we assume perfect competition throughout, the resulting investment and decommissioning decisions are formally equivalent to those that would emerge with a sufficiently large number of identical agents.

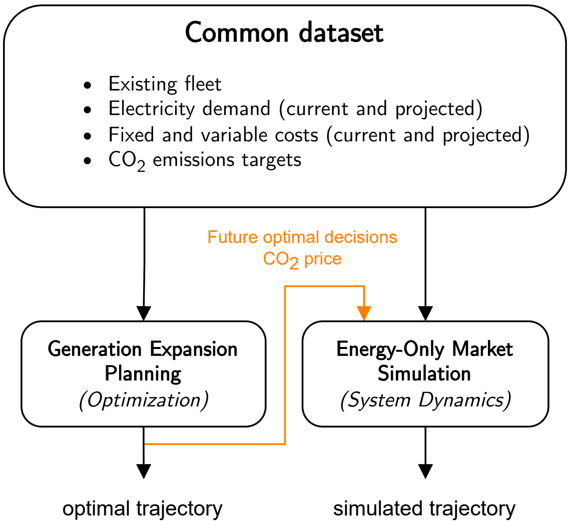

Figure 1: Modeling framework overview

Note: The linkage between the GEP and SD models represented by the orange arrow is optional.

developments are anticipated and impact future price formation – and in turn govern investment decisions – thus exemplifying an issue of endogeneity we will come back to in Section 2.2.

This is the main reason why we complement our modeling framework with a GEP model. That is, we feed the SD model with two relevant outcomes from the GEP model (see Figure 1). First, some information about future capacity developments from the GEP model can be used in the SD model when the representative investor forms future price and revenue anticipations (see Section 2.2). Second, because we do not explicitly model the market for carbon emissions, the shadow price associated with the annual constraint on emissions from the GEP model is used as an exogenous carbon price signal in the SD model (see Section 3). Finally and intuitively, the GEP model also constitutes a valuable normative benchmark for optimal capacity developments against which we will assess the (deviations in) outcomes resulting from the SD model.[10]

The GEP model is a standard constrained pluriannual cost-minimization problem.[11] In short, its objective function is the expected net present value of total systems costs (operating costs + investment costs + cost of rationing non price-responsive consumers at the value of lost load). Decision variables include short-term operations and long-term entries and exits. Several constraints govern the market clearing, various generation and storage asset operations, asset fleet dynamics, and carbon emissions. The different technologies we represent are described in Section 3. We now turn to the description of the SD model.

2.2 System Dynamics market model

The SD market model simulates the evolution of the generation mix as the result of successive market outcomes and stepwise investment and decommissioning decisions over time. Compared to

[10]This alleviates the somewhat arbitrary nature of the reference point often used in simulations, and it establishes a bridge between the long-term market design simulation literature and the prospective analysis literature.

[11]Because the formulation of this problem is standard, its detailed presentation is relegated to Appendix B.

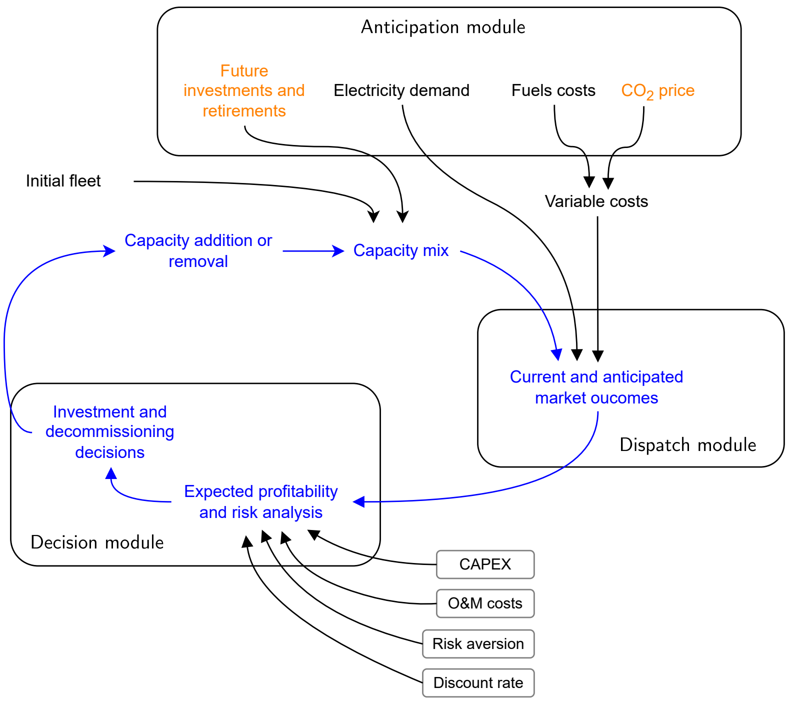

Figure 2: Causal-loop diagram of the SD market model

Note: Elements in blue are endogenously determined and updated at each iteration of the loop delineated by blue arrows. Inputs in orange can be taken from the GEP model, whereas those in black are exogenously calibrated.

intertemporal optimization or equilibrium problems that are generally processed by a numerical solver computing optimal values for all current and future decision variables simultaneously, this approach explicitly unpacks both the agents' decision-making process and the dynamics of capacity development. While this has the potential to capture important behavioral and transitory effects (e.g., path dependency), it also entails that we have to explicitly address the issue of endogeneity (i.e., circularity or co-determination) between current and future decisions, that is the formation and adjustment of (long-term) anticipations.

Specifically, individual investment and decommissioning decisions are determined endogenously and sequentially each year of the simulation period on the basis of an economic profitability assessment for a given set of investor behavior assumptions and range of future market outcomes. The SD model is composed of three interconnected modules to enable this assessment: First, an anticipation module that produces reference scenarios for market fundamentals over time. Second, a dispatch module that generates market prices as well as generation and storage patterns in future years for given reference scenarios. Third, a decision module that makes annual investment and retirement decisions using information about costs and revenues from the first two modules. The outcome of the decision module is then fed into the anticipation module, adjusting reference scenarios and initiating an iterative loop. The three modules and their linkages are graphically shown in the causal-loop diagram in Figure 2. We describe them in turn below.

2.2.1 Anticipation module

The anticipation module produces reference scenarios on the basis of two types of long-term market fundamentals that govern current investment and decommissioning decisions. First, some parameters such as future demand, fuel prices, and carbon prices are set and calibrated exogenously (see Section 3.1). Second, anticipations must also be formed on future endogenous variables (i.e., future investment and decommissioning decisions that define the generation mix in the long term) which affect investment and decommissioning decisions today, and vice versa. This section focuses on the formation of the second type of anticipations over a 'prospective horizon'; namely, the time period over which investment projects assessed in the current year of the simulation operate and are remunerated (and equivalently for retirement projects).[12]

As a (rational) way of dealing with deep uncertainty and cognitive limitations associated with their long-term decision-making, investors may use heuristics or rules of thumb to alleviate associated computational complexity and informational requirements (e.g., Simon, 1955; Baumol and Quandt, 1964). For instance, heuristics can be utilized to forecast future relevant factors (Brock and Hommes, 1997), such as backward-looking adaptative expectations (Cepeda and Finon, 2011). Alternatively, simplified models considering only variables of first-order importance can be built (Gabaix, 2014), information that is costly to obtain and process can be ignored (Reis, 2006), and planning horizons can be truncated and sliding (Quemin and Trotignon, 2021). Moreover, limited sophistication in forecasting future entries and exits over the whole prospective horizon can be justified by the fact that anticipating other agents' decisions is complex, especially without complete long-term markets that may allow for the coordination of agents towards the first-best outcome (e.g., Williamson, 1975; Van Huyck et al., 1990; Felder, 2002).

For expositional clarity, we consider two polar cases of anticipation sophistication labeled 'static' and 'dynamic' anticipations. This notably allows us to avoid using in-between ad-hoc anticipation heuristics or rules that would make our results dependent on arbitrary modeling choices. The two cases are graphically illustrated in Figure 3 for a stylized example of a fossil technology phaseout.

Static anticipation. The first and simplest case consists in not considering any future decisions throughout the prospective horizon. That is, at a given point in time, the existing fleet is maintained online with no new entries until existing assets reach the end of their lifespans and retire without early economic exits. This simplifying assumption of myopia, albeit somewhat extreme, is often (implicitly) made in the related literature (e.g., de Vries et al., 2013; Chen et al., 2018). In Figure 3, in any given year of the simulation (here 2025), before any investment or decommissioning decision is made on that year, installed capacity 0 inherited from the previous year is prolonged into the future to form the anticipated trajectory A. Note that this trajectory is constant as long

[12] The model features developments to represent build times that are not used in this article and left for further work. Here, build times are neglected and assets start producing instantaneously when the decision is made.

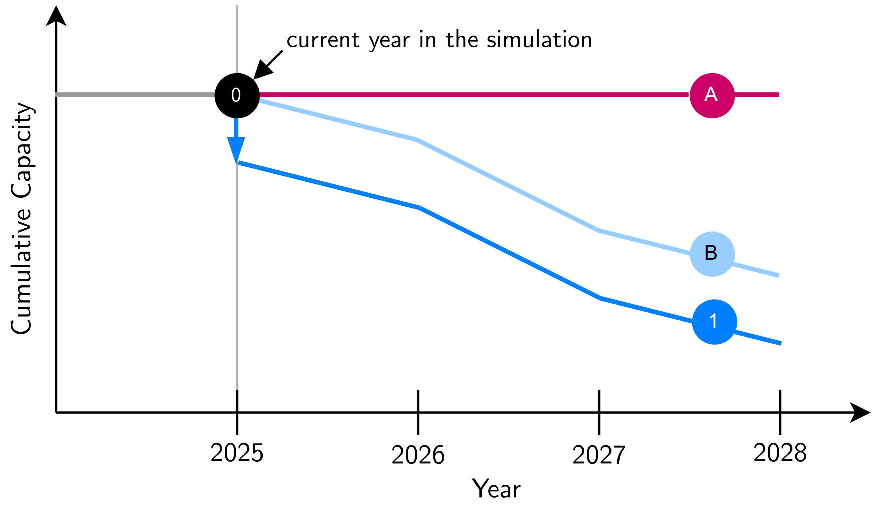

Figure 3: Illustration of static vs. dynamic anticipation of future decommissioning decisions

as no unit reaches the end of its lifespan (as is the case here until 2028).

Dynamic anticipation. The second case introduces sophistication in anticipating the evolution of the generation mix over time. In line with standard prospective analysis methodology utilized in practice to inform investment and retirement decisions, the GEP model is leveraged to provide an optimal capacity development plan.[13] Specifically, in a given year, the GEP model is run from the current state of the fleet, yielding optimal capacity trajectories by technology until the end of the prospective horizon. Then, the anticipated evolution of the generation mix is defined as per the optimal yearly capacity changes beyond the current year, while entry and exit decisions for the current year are left for the decision module to determine. This guarantees that in the benchmark case (i.e., assuming perfect information, rationality and risk neutrality) the decision module arrives at the optimal decisions for the current year that are congruent with the entire optimal capacity development plan derived from the GEP.[14] This also constitutes a conservative assumption when assessing the deviations induced by moving away from the benchmark case (Section 4). In Figure 3, the GEP model run from the initial conditions 0 yields the full optimal trajectory 1, and the dynamic anticipation of future entries and exits B is obtained by only keeping optimal decisions beyond the current year.

Moreover, the anticipation module is designed to accommodate two types of deviations from the above perfect dynamic anticipation case. First, we can relax the perfect information assumption and introduce various biases in the anticipation of future fundamentals, regarding either the

[13]In practice, firms and investors can carry out such prospective analysis in-house or have it provided by external consultancies. In most cases if not all, this relies on GEP-style optimization tools (e.g., PROMOD by Hitachi Energy or PLEXOS by Energy Exemplar).

[14]This convergence result is verified in the simulations (see Section 4.1) and arises by construction: in the spirit of a rational expectations equilibrium framework, we seek the fixed point between equilibrium and optimal beliefs about future capacity expansion. Importantly, note that this result also requires the anticipation of the associated future prices and revenues over the entire lifespan of all assets (see Section 2.2.3 and footnote 18).

parameters used as inputs in the GEP model (e.g., demand level, technology costs) or its outputs (e.g., carbon price, capacity trajectories). The introduction and choice of these biases is arbitrary, but this should be seen as a complement to our analysis that resorts to two polar anticipation cases. Second, we can choose the frequency at which the GEP model is called to update the anticipation of future entries and exits. While annual updates would be ideal, in practice the associated cost and complexity of this exercise may justify an update every couple of years or more.[15]

2.2.2 Dispatch module

The SD model embeds an economic dispatch module to represent short-term operations and market prices. This type of linkage, first introduced in Dyner et al. (2011), is preferred over the 'revenue curve' (e.g., Ousman Abani et al., 2018) and 'scarcity rents curve' (e.g., Kraan et al., 2019) approaches that exogenously define a direct relation between the level of capacity and assets' revenues. Although computationally heavier, our approach allows for a more accurate representation of key features of decarbonized power systems (notably the time variations of weather-dependent generation and the dynamics of storage), and consequently of the price and revenue distributions. Below we first describe the module's structure and main assumptions, and then specify when and how it is utilized within the SD model workflow.

The dispatch model is formulated as a standard short-term cost-minimization problem whose objective function is the total operating cost; i.e., the sum of variable costs and cost of rationing price-inelastic consumers (set at the VoLL). We consider an hourly resolution and to alleviate the computational burden, we divide the annual problem into sub-problems with a rolling horizon and a 'look-ahead interval' (i.e., a final or continuation period that is included in each sub-problem whose solutions are discarded in the current sub-problem but utilized in the subsequent sub-problem). We set these parameters to 1 month and 24 hours, respectively, and storage assets are dispatched over these optimization steps assuming perfect foresight.

Implicitly, this representation of short-term operations assumes that the sequence of short-term markets is frictionless and yields an optimal outcome.[16] This important assumption is deliberate, as we wish to zero in on the long-term aspects of a canonical energy-only market that would still prevail if current short-term markets were improved through more integration and finer granularity. For the same reason as well as for consistency with the case study presented in Section 3, we keep the dispatch model as parsimonious as possible. Although this version of the dispatch model does not represent short-term uncertainties, seasonal storage, ancillary services and grid congestion, its generic and modular implementation is amenable to such developments and refinements.

Finally note that the dispatch module is called and run in two different places in the SD model:

[15] Given the relative monotonicity of our case study (see Section 3) and the low degree of stochasticity, the optimal capacity trajectories defining the dynamic anticipation are only computed once at the beginning of the simulation period in this version of the model. This has negligible impacts on our results while reducing computational burden.

[16] This notably implies the absence of market power and short-term non-convexities.

First, it is primarily utilized to convert long-term fundamentals from the anticipation module into future market prices and revenues as well as generation, storage charging and renewable curtailment patterns. This dispatch is dubbed "*virtual dispatch*" (Tao et al., 2021) as it is run for future dates on the basis of different anticipations of the future state of the system. Second, once investment and decommissioning decisions are made in a given year, the dispatch module is run to simulate the (actual, not virtual) short-term market outcomes for that year, before moving on to the following year. This final run is notably used as a basis to compute different metrics in Section 4.

2.2.3 Decision module

The decision module consists of a loop that considers all possible investment and retirement decisions and iteratively selects the most profitable one at each step in the simulation until none is left. Specifically, economic profitability is assessed on an annual basis from a representative investor perspective using a Net Present Value (NPV) criterion – possibly adjusted for risk aversion. The NPV approach is a well-established tool in the literature to appraise and compare assets available for investment and retirement.[17] The representative investor perspective is at the core of the SD approach and has been motivated above (Section 2.1).

Decision criteria and loop. The NPV associated with each decision is computed using relevant costs (i.e., avoidable costs) and market revenues over a certain time horizon that all depend on the nature of the decision or the underlying asset (e.g., investment or closure, existing or new asset).

- Costs: We consider two types of fixed costs, the investment cost (CAPEX) and fixed Operations and Maintenance (O&M) costs. CAPEX is a sunk cost that cannot be recovered when an existing unit is retired. By contrast, fixed O&M cost is due when the unit is in operation but can be saved by decommissioning it. Some technologies also have a variable operating cost, that is fuel and carbon costs for thermal assets or charging cost for storage assets.

- Revenues: Annual revenues accruing to all assets over their entire lifespans are computed using the hourly virtual short-term dispatch module (Section 2.2.2) with the long-term fundamentals from the anticipation module (Section 2.2.1) as inputs.[18] What matters is the stream of net revenues, that is short-term market revenues minus variable costs.

- Horizon: The profitability of a potential investment in a given year must be assessed over its whole lifetime. By contrast, the timing of a potential retirement decision is more complex, as

[17] Although there is an option value in deferring decisions to invest in new assets under uncertainty and investment irreversibility (e.g., Dixit and Pindyck, 1994; Rios-Festner et al., 2019), our approach makes no arbitrage between investing now or a few years later (i.e., no real-options valuation). Yet we implement a procedure that reflects some optionality for decommissioning decisions. Our approach does also not consider portfolio synergies across assets.

[18] It is by now clear that there are two layers of foresight in the model – one for future entries and exits, another for future prices and revenues. Assuming that the forward-looking anticipation of future inframarginal rents is not truncated ensures that current investment and retirement decisions are optimal from a system perspective in the case of perfect dynamic anticipation of future entries and exits (see Section 2.2.1 and footnote 14).

temporary losses in the short term can be offset by larger gains in the long term.[19] To capture this, we implement a simplified procedure whereby retirement occurs only if revenues do not cover fixed O&M costs both in the current year and over the asset's remaining lifespan.

The NPVs of all potential decisions in a given year are computed as described above and the investor picks up the one with the highest NPV per megawatt of capacity in absolute terms. The decision-making process thus places investment and retirement decisions on an equal footing and treats both types of decisions simultaneously.[20] Once a decision is made, the asset fleet is adjusted for the corresponding capacity addition or withdrawal, which affects the economic profitability of all other units – be they installed or under consideration. That is, at each iteration of the loop, expected market revenues and NPVs of all assets are updated on the basis of the iterative evolution of the asset fleet. Importantly, the profitability of all previous decisions made in the current year is reassessed at each iteration: if an earlier decision becomes unprofitable because of some following decisions, it is called off; and only those decisions that stay profitable until the end of the iterative loop become firm and effectively materialize.

Additionally, note that we implement a standard modeling artifact to account for those years in the profitability assessment that extend beyond the simulation period. Specifically, we assume market revenues earned in the last year of the simulation period are duplicated and repeated over the following years until the entire asset lifetime is covered in the profitability assessment. Because the 'edge effects' induced by this artifact become increasingly prevalent as the end of the simulation period nears, our interpretation of the simulation results in Section 4 will essentially focus on the time window where they are less distorted (i.e., in the first part of the simulation period).

Finally, the iterative loop in a given year terminates when one of the two following conditions is met: either there is no profitable decision left, or a given state of the asset fleet (i.e., the number of units per technology, which is stored at the end of each iteration) is reached for the second time.[21] An algorithmic description of the decision loop is provided in Appendix C.

Uncertainty and risk preferences. We adopt a simple approach to introducing uncertainty around the reference scenario produced by the long-term anticipation module (Section 2.2.1). We follow Neuhoff et al. (2022) and bypass the explicit modeling of multifaceted uncertainty, e.g. on demand, commodity prices, regulatory changes and so forth.[22] That is, we consider that risk bears directly on the discounted sum of net revenues for a given asset in the reference scenario, denoted $\bar{r}$. Specifically, our main case considers that asset-specific net revenues are uniformly distributed

[19] For simplicity, mothballing decisions are not considered. See Ousman Abani (2019) for these developments.

[20] The related literature typically focuses on investment and retirement issues in isolation, often effectively modeling only one. When both decisions are endogenous, they are typically modeled sequentially, which is not inconsequential for market outcomes. Our framework circumvents and goes beyond these issues.

[21] This preempts infinite back and forth and addresses indeterminacy due to unit indivisibility (see Section 4.1).

[22] This single risk can be thought of as the aggregate of all risks, but note their cross-effects are not captured.

between 0 and $2\bar{r}$.[23] In contrast to Neuhoff et al. (2022), therefore, the central value around which we cast our probability distribution is endogenously determined in the model.[24]

In the face of uncertainty, we consider that the representative investor can exhibit different degrees of risk aversion, including risk neutrality. There are many reasons why investors and firms (or firms acting on behalf of investors) de facto behave as if they were risk averse, resulting in a higher utility from more stable profits. These include, inter alia, hedging demand, corporate risk management policies (e.g., financial and operational constraints) or costs associated with financial distress (e.g., Froot et al., 1993; Bessembinder and Lemmon, 2002; Willems and Morbee, 2010; Acharya et al., 2013; Jagannathan et al., 2016). Additionally, in the electricity industry where assets are often capital-intensive with long lifetime, investment decisions are the result of careful profitability assessments, and one may intuitively expect risk aversion to prevail (e.g., Vázquez et al., 2002; Neuhoff and de Vries, 2004; Abada et al., 2019).

There are different approaches to representing risk aversion, including risk-adjusted discount factors (e.g., in the spirit of the CAPM), coherent risk measures (e.g., a linear mixture of expected surplus value and conditional value-at-risk), and concave utility functions. Although the first two approaches allow for a more detailed and state-of-the-art analysis of risk impacts, they deploy a heavier machinery than the third approach that is irrelevant given our rather crude representation of uncertainty. We thus assume that the investor preference for stable and secure profits is described by a concave von Neumann-Morgenstern utility function $\mathcal{U}$. Additionally, as is standard in the related literature (e.g., Petitet et al., 2017; Fraunholz et al., 2023), we consider that risk aversion applies directly on the (distribution of the) asset-specific discounted sum of net revenues.[25]

We choose a functional form that satisfies the property of constant relative risk aversion. This property ensures that the coefficient of risk aversion does not vary with the economic value of the decision under consideration, which typically is of a different order of magnitude for investments and retirements. Specifically, following Petitet et al. (2017), $\mathcal{U}$ is defined by

$$\mathcal{U}(\mathbf{r}) = \begin{cases} 1 - \exp\left(-\alpha \mathbf{r}/\bar{r}\right) & \text{for} \quad \alpha > 0 \\ \mathbf{r} & \text{for} \quad \alpha = 0 \end{cases}$$

with α the coefficient of (constant relative) risk aversion, $\mathbf{r}$ the random discounted sum of net revenues, and $\bar{r} = \mathbb{E}\{\mathbf{r}\}$ by definition.[26] Under risk neutrality ($\alpha = 0$), the investor considers the mean of the net revenues distribution $\mathbb{E}\{\mathcal{U}(\mathbf{r})\} = \mathbb{E}\{\mathbf{r}\} = \bar{r}$ when computing the NPV of a given

[23]In Appendix E, we illustrate how our qualitative results are unchanged when we vary the interval of the uniform distribution or when we consider another distribution type for the random variable.

[24]Neuhoff et al. (2022) consider the last year for which liquid futures contracts exist and assume constant electricity prices and market values.

[25]This assumption is reasonable for our purposes, and it also allows us to keep the model tractable given the other key modeling details and specificities that we need to account for. Specifically, we do not develop a recursive utility model à la Kreps-Porteus or Epstein-Zin as the issue of intertemporal substitution is of second-order consideration and applicability for the problem at hand. There is thus no need to disentangle risk aversion from intertemporal substitution, and we can apply risk averse preferences directly on the overall discounted net revenue streams.

[26]One can easily check that constant relative risk aversion holds, that is $-\bar{r}\mathcal{U}''(\cdot)/\mathcal{U}'(\cdot) = \alpha$.

asset. Under risk aversion ($\alpha > 0$), the investor considers the certainty equivalent of net revenues $r^\star$, that is the certain revenue that yields the same utility as the expected utility over the random distribution of revenues, $\mathcal{U}(r^\star) \equiv \mathbb{E}\{\mathcal{U}(\mathbf{r})\}$. The certainty equivalent $r^\star$ is decreasing with α and tends to $\bar{r}$ in the limit as α goes to zero. Analytical details are relegated to Appendix E.

3 Case study

In this section, we introduce and document our Californian case study. We first describe the fundamentals used to calibrate the model (Section 3.1) and then present the optimal simulation results derived from the GEP model (Section 3.2). All input data are available here: Zenodo/Dataset.

Before we proceed, a short discussion on market design is in order. In practice, the Californian system is composed of a nodal electricity market with a soft offer cap at 1,000 \$/MWh, a mandatory resource adequacy requirement with no formal capacity market, an emissions trading system (ETS), and a renewable portfolio standard (RPS) program. However, we intend to leverage our stylized case study to represent a canonical energy-only market (EOM). Bearing in mind that our study is for illustrative purposes, we thus make several simplifications in terms of implementation.

We consider a zonal market with an hourly resolution over a 20-year period, focusing only on wholesale electricity (ancillary services are outside the scope of this paper). There is no offer price cap and the single hourly price can go up to the VoLL set at 15,000 \$/MWh. Moreover, we model an isolated system (no interconnection) and do not represent the internal network ('copper plate' assumption).[27] Although we do not formally account for the resource adequacy requirement, we set load scenarios in line with the 'one-in-ten' regulatory criterion. Similarly, we do not represent the RPS program, but note that CPUC (2019) found the RPS constraint to be non-binding (the associated shadow price is zero) and decarbonization to be driven only by the constraint on emissions. Likewise, the ETS is not explicitly modeled and the price of carbon is set as the shadow price associated with annual emission targets. In sum, this stylized setup allows us to capture the essence of EOM outcomes and to purposely assess how they hinge on investor behavior assumptions.[28]

3.1 Calibration

We use and adapt data from three open-data sources: the integrated resource planning (IRP) exercise by the California Public Utilities Commission (CPUC, 2019), Ninja Renewables (Ninja Renewables, 2021), and historical data from the California Independent System Operator (CAISO, 2018, 2019). Below, we describe how we calibrate demand, supply and decarbonization parameters for both the optimization (GEP) and simulation (SD) models.

[27] Imports and exports are relatively small compared to domestic generation and consumption: imports cover 17% of total supply today (CAISO, 2022) and decline in CPUC's scenario to a slight net exporter situation in the 2030s.

[28] See Bruninx et al. (2020) and Osorio et al. (2021) for related modeling approaches that jointly represent these different markets. Although they also look into investment decisions in the electricity sector, their focus is on policy design assessment and interactions, and they assume perfectly rational agents.

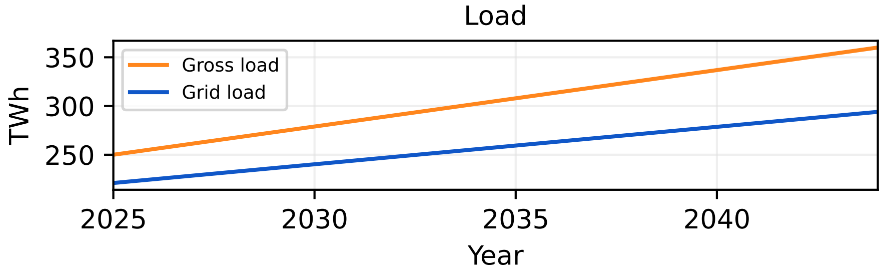

Figure 4: Load assumptions (grid load = gross load – distributed solar generation)

Demand. We consider that demand is fully inelastic to price variations and does not adjust to installed capacity. While this simplifies the model and its calibration, note that this is innocuous for the EOM performance we aim to appraise. Indeed, price-inelastic demand does not fundamentally alter the peak-load pricing logic inherent to the EOM (e.g., Joskow and Tirole, 2007; Joskow and Léautier, 2021).[29] Specifically, gross load is exogenously given and assumed to increase linearly between two given data points taken from CPUC, namely 250 TWh in 2025 and 360 TWh in 2045. Similarly, we assume that distributed solar generation is increasing linearly between two given data points, 29 TWh in 2025 and 66 TWh in 2045, and we subtract it (on an hourly basis, see below) from the gross load to obtain the grid load. The result is graphically depicted in Figure 4.

Next, we convert the above annual amounts into hourly time series. To do so, we start from two historical years, 2018 and 2019, for which we obtain hourly load from CAISO (CAISO, 2018, 2019) and hourly capacity factors for wind and solar from Ninja Renewables (Ninja Renewables, 2021). These constitute our two representative scenarios capturing 'short-term' uncertainties (i.e., which resolve as real time nears). Working with historical data allows us to infer and then utilize realistic correlations between various sources of uncertainty such as load and weather. We then set the hourly load profile so as to reflect the 'one-in-ten' capacity adequacy criterion that applies in California. We proceed in four steps: First, we normalize both series. Second, we set 2018 as the representative year for the one-in-ten peak and 2019 as the representative average year. Third, we scale the normalized 2018 profile homothetically to have a demand peak that is 15% larger than that in 2019. Fourth, we scale the two load profiles so that they sum up to the annual amounts in Figure 4, considering 10% and 90% probabilities for 2018 and 2019 respectively.

Supply. There are two types of technologies: 'exogenous' technologies, whose installed capacities are exogenously given and in line with planned evolution over time (e.g., by mandate or regulation), and 'endogenous' technologies for which entry and exit decisions are explicitly modeled.

The set of 'exogenous' technologies consists of Combined Heat and Power (CHP), nuclear, existing and planned wind and solar (with shorthand 'E&P'), as well as geothermal, biomass and

[29] In other words, demand elasticity is not a theoretical requirement for the EOM to yield optimal outcomes. With price-inelastic demand, demand is curtailed when it exceeds capacity and the price is set at the VoLL.

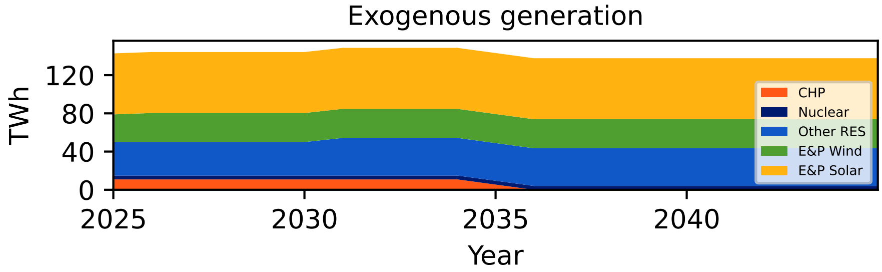

Figure 5: Annual generation of 'exogenous' technologies

small hydro, which we group into one category (labeled 'Other RES').[30] Hourly availability factors for these technologies are set on the basis of data from CPUC (2019), which we translate into an hourly resolution using Ninja Renewables (2021) if necessary. Figure 5 depicts the evolution of generation volumes over time for these technologies.

The set of 'endogenous' technologies is further divided in two groups on the basis of decisions available to the representative investor: existing fossil-fired dispatchable technologies – Peaker and Combined Cycle Gas Turbine (CCGT) – can only be decommissioned (i.e., no new investment is possible), whereas Solar and Storage technologies can also be invested into. The storage technology we consider is generic and has the characteristics of a lithium-ion 6-hour Battery Energy Storage System (BESS) with a 85% roundtrip efficiency. The technical and economic parameters of these four technologies are given in Table 1.

We make several realistic or mild assumptions to streamline the simulations and their interpretations. That is, they allow us to isolate the impacts of varying investor behavior assumptions and are largely inconsequential for the qualitative nature of our results. First, all 'endogenous' technologies have a common weighted average cost of capital of 8%.[31] Second, they have a common lifespan of 25 years that is longer than the simulation horizon. Third, investments are realized with no build time (i.e., new capacities are built and start operations on the year the investment decision is made). Fourth, in our central scenario, all technologies have a common indivisible unit size of 500 MW. Taken together, these four assumptions essentially guarantee that investment and decommissioning decisions are not structurally tilted towards specific technologies. Taken individually, each assumption is mild and simplifies the anticipation and decision modules introduced in Section 2.2.[32] Although the first three assumptions have relatively straightforward and innocuous implications (e.g., the higher the WACC, the lower the investment volume), Section 4.1 provides a sensitivity analysis of the fourth assumption as a basis for a general discussion of capacity unit discreteness/indivisibility for modeling outcomes.

Finally, we set the initial conditions – that is, the capacities installed in 2025 – so as to start the

[30]Solar technology on the supply side corresponds to grid-scale solar PV (recall that distributed solar generation is accounted for exogenously on the demand side).

[31]For a literature review on the representation of the cost of capital in energy system models and a discussion of the associated impacts on model outcomes, see Lonergan et al. (2023).

[32]For instance, the second assumption implies that we do not have to address the issues of refurbishing, repowering or closing the assets that are built during the simulation period.

Technology	Available decision	CAPEX [USD/kW-yr]	Fixed O&M [USD/kW-yr]	Variable cost [USD/MWh]	Carbon intensity [$kgCO_2$/MWh]
Peaker	D only	42.5	20	51*	610
CCGT	D only	117	30	31*	370
Solar	I & D	72	9	0	0
Storage	I & D	108	13	-	-

Table 1: Technical and economic parameters for 'endogenous' technologies

Note: The letters D and I refer to decommissioning and investment, respectively. The superscript * indicates average values over the simulation period and the symbol - denotes parameters whose measurement is not straightforward due to yield and intertemporal use issues. For simplicity, Solar's variable cost and carbon intensity are assumed to be zero. In addition, all technologies are assumed to have no build time, the same weighted average cost of capital (8%), and the same lifespan (25 years) that extends beyond the simulation duration.

simulations from an equilibrium state that is congruent with the anticipation and decision modules described in Section 2.2. Specifically, CCGT, Peaker and Storage capacities are determined by running the GEP model for the year 2025 alone – they amount to 11, 19 and 10 GW, respectively. 'Endogenous' Solar capacities for 2025 are initialized to zero since the existing and planned (E&P) fleet is already accounted for in the set of 'exogenous' technologies.

Decarbonization. We consider that the climate change externality associated with electricity production is fully internalized through a carbon price signal that is commensurate with the stringency of decarbonization targets.[33] In our setup, the shadow price associated with the constraint on CO_2 emissions is equivalent to the equilibrium price that would arise in a perfectly competitive market for permits (Montgomery, 1972).[34] Specifically, we assume that the annual cap on CO_2 emissions is decreasing linearly over time and we calibrate it with a linear interpolation between 2025 and 2045 emission levels. We compute the former as the emissions resulting from the initial brownfield fleet (31 $MtCO_2$) while we get the latter from CPUC (12 $MtCO_2$). Figure 7 depicts the resulting emissions cap trajectory (left panel) and associated optimal price signal (right panel).

3.2 GEP results (optimal trajectories)

We run the GEP model with continuous capacity adjustment to determine the optimal capacity trajectories for 'endogenous' technologies (Figures 6–7). Regarding new developments, Solar and Storage reach 84 and 59 GW of installed capacity in 2045 respectively with a quasi-linear trend.[35] Regarding existing assets, Peaker capacity is reduced by 6 GW (down to 13 GW), whereas CCGT capacity remains unchanged. The annual emissions constraint is satisfied and binding every year with the shadow price of carbon rising over time (399 $/$tCO_2$ in 2045).[36]

[33] As a result, the failure to decarbonize in line with targets identified in Section 4.2 cannot be attributable to an inefficient carbon price signal – rather, this has to do with investor behavior assumptions. See Ruhnau et al. (2022) for a recent review and cross-model comparison of carbon price impacts in electricity market models.

[34] Since we do not explicitly model the ETS, we abstract away from inter-annual flexibility via banking, which may induce second-order changes in emission and price levels (e.g., Rubin, 1996; Schennach, 2000); and price containment mechanisms, which may have an impact on investment decisions (e.g., Burtraw et al., 2022; Cason et al., 2023).

[35] This trend is essentially driven by the input data and foreseen evolution of load over time (Figure 4).

[36] As a quick sanity check, we compare our results to those of CPUC: Solar and Storage capacities amount to 64.3 and 50.8 GW in 2045 respectively; Peaker and CCGT capacities are reduced by 4.2 and 1.8 GW over 2025–45, down

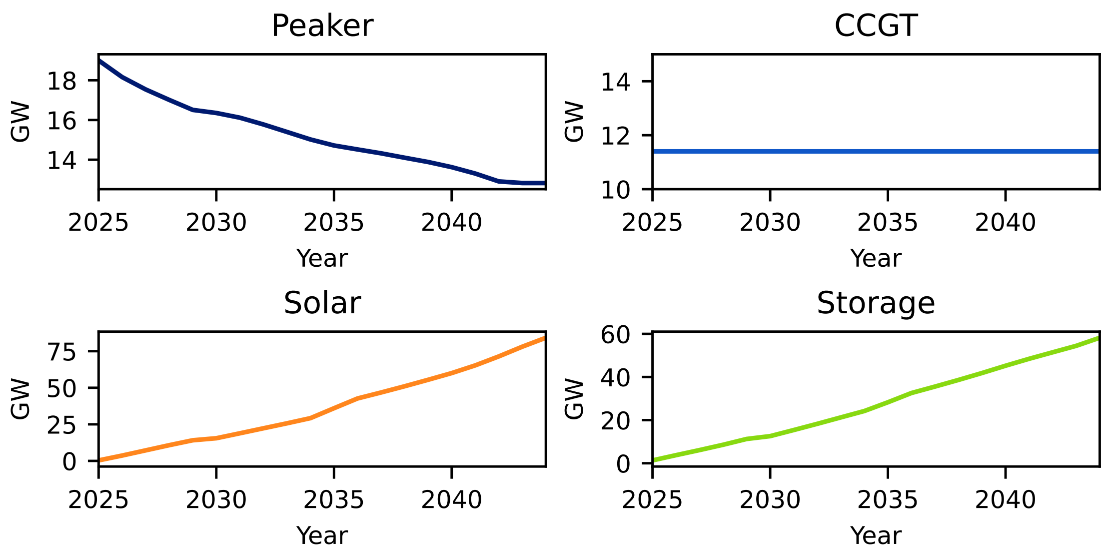

Figure 6: Optimal capacity trajectories for endogenous technologies (GEP results)

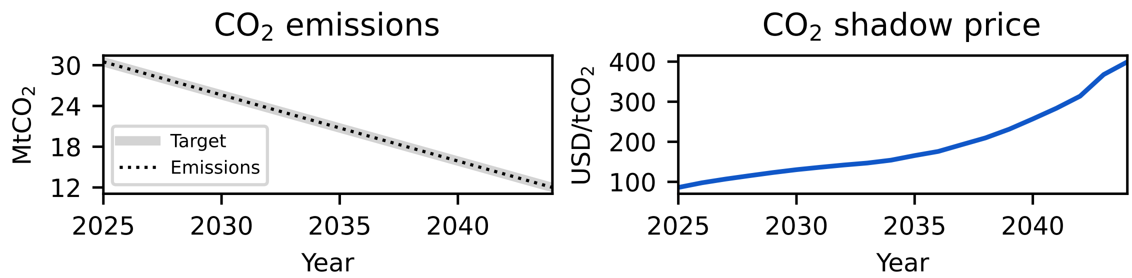

Figure 7: Annual CO_2 targets and emissions (left) and shadow price (right) (GEP results)

Interestingly, note that only a few hours of load rationing occur in the first years and that they vanish afterwards, even as decarbonization progresses (Figure 9, dashed line).[37] This result holds in our case study because of the massive build-out in both new Solar and Storage capacities which is driven by decarbonization targets and, as a by-product, satisfies the system's capacity needs. That is, the emissions constraint always binds and prevails over the overall capacity constraint.

One may rightly wonder how Peakers are able to recover their costs in such conditions, i.e. in the absence of inelastic load rationing events with prices reaching the VoLL. One must first recognize that the formation of electricity prices (i.e., the system's hourly marginal values of electricity) in carbon-constrained systems with high shares of renewables and storage is complex and "*determined dynamically by demand and intertemporal storage decisions, breaking the static logic of 'merit order' with dispatchable generation*" (Ekholm and Virasjoki, 2020). In particular, prices can settle above the highest conventional generator's variable costs because of storage's roundtrip efficiency and intertemporal arbitrage, possibly forming "*price plateaus*".[38] Moreover, prices can factor in long-term cost components when capacity additions and retirements are endogenous in the model (e.g.,

from 8.6 and 16.2 GW respectively; and the carbon price is found to reach 403 $/tCO_2$ in 2045. Although there are some quantitative deviations, our modeling results for a stylized system are qualitatively very comparable overall.

[37] Again, this is in line with CPUC's results where the shadow price associated with the Reserve Margin constraint is 63 $/kW-yr in 2026, 0 $/kW-yr in 2030 (i.e., the constraint is not binding) and 1 $/kW-yr in 2045.

[38] Recall that storage units are optimized over a one-month window with perfect foresight.

Mallapragada et al., 2021). In this context, all conventional generators (including Peakers) pocket sufficient inframarginal rents to recover their fixed costs.

Running the GEP model allows us to illustrate numerically how, in line with theoretical principles, cost recovery is ensured for all assets with no 'excessive' rents. Specifically, we compute the Cost Recovery Ratio (CRR), i.e. the ratio of net market revenues (price – variable costs) to fixed investment and operational costs.[39] GEP column in Table 2 shows that each endogenous technology recovers exactly 100% of its costs. Note that for initially existing (brownfield) assets, economic viability only requires fixed O&M costs to be recovered (respectively 20 and 30 USD/kW-yr for peaker and CCGT) as CAPEX is sunk. This explains why existing Peakers recover only 32% of their fixed costs, which corresponds to the share of their fixed OPEX. Since Peakers are at the margin regarding total capacity – i.e., Peaker capacity is adjusted downward against new capacity additions –, retained units just break even, exactly recouping their fixed OPEX with no extra rent to cover their CAPEX. By contrast, infra-marginal CCGTs recover 91% of their total fixed costs, which is sufficient to cover the 20% share of fixed OPEX but also recoup a certain amount of CAPEX though not its entirety.

4 Results and discussion

In this section, we present our simulation exercise and discuss the results. First, we examine the conditions under which the SD energy-only market (EOM) model is able to replicate the optimum as defined in Section 3.2, and we discuss the impacts of unit indivisibility. Second, we explore how the simulated EOM outcomes deviate from the optimum when we relax these conditions.

Our quantitative analysis is based on the following indicators:

- **Capacity trajectories** to assess how investments and retirements for each technology compare with the optimal ones over time,
- **Carbon emissions** to check if decarbonization targets are achieved,
- **Total cost** as an overall cost efficiency indicator,
- **Cost recovery ratios** (CRR) as technology-specific capacity remuneration indicators,
- **Average marginal cost**, interpreted as an average price, as an affordability indicator,
- **Loss of load expectation** (LOLE) as a system-wide capacity adequacy indicator.

[39]Formally, with the notations of Appendix A, for a given vintage of a technology t invested in year y, one has

$$\mathrm{CRR}_{t,y} = \sum_{k=y}^{\min(\#\mathcal{Y},\mathcal{L}_u)} \beta^{k-y} \frac{\sum_{w\in\mathcal{W}} \Pi_w \left[\sum_{h\in\mathcal{H}} \left[\frac{q_{t,k,h}}{n_{k,t}k_t} \left[\lambda_{k,h} - VC_{k,t}\right]\right]\right]}{IC_{y,t} + OC_{y,t}}.$$

Table 2 shows the CRR averaged over vintages, but note that $\mathrm{CRR}_{t,y}$=100% holds for all vintages y for endogenous technologies. In Section 4.2, we compute and analyze $\mathrm{CRR}_{t,y}$ across vintages.

Since some simulation or sensitivity cases are computationally demanding (notably as we reduce the capacity step size, see Section 4.1.2) we run the simulations and present the results only over the 2025–35 horizon. Yet note that this is largely innocuous for the validity of our analysis beyond this window given the relative 'monotonicity' of our case study, as evidenced by the quasi-linear trends for capacity trajectories in the optimal case over the 2025–45 horizon (Figure 6).

4.1 EOM outcomes with idealistic assumptions

4.1.1 Definition of idealistic assumptions

As discussed in the Introduction, it is often regarded as well-established that the long-run optimum can in principle be decentralized through competitive market prices. Although the prerequisites are regularly stated in general and concise terms (e.g., 'perfect markets', 'Arrow-Debreu economy'), some authors including Rodilla and Batlle (2012), Newbery (2018) and Joskow (2008, 2022) establish more detailed lists that we transcribe below:[40]

1. Agents (i.e., buyers and sellers) are anonymous, atomistic and fully rational (i.e, price-taking and non-strategic behavior);
2. Agents have convex cost and utility functions (no non-convexities, no economies of scale);
3. Capacity, generation and consumption levels can be adjusted continuously (no lumpiness);
4. There is perfect information with well-informed agents and no asymmetries;
5. There is a complete set of markets covering all relevant contingencies and over all relevant timescales, including markets for insurance.

Let us translate these canonical assumptions to our modeling framework. The first is satisfied by construction since we consider one representative investor who behaves non-strategically and makes investment and retirement decisions on the basis of a competitive profitability assessment. The second is met by ruling out cost non-convexities and because we do not represent the demand side formally. The third does not hold since capacity increments and decrements are discrete by implementation in the SD model. As capacity indivisibility also holds in practice, its impacts on market outcomes are analyzed and discussed below (Section 4.1.2).

Next, because the decision-making process crucially hinges on available information, we break down the fourth assumption into two sub-assumptions:

A1. Agents have perfect information on all exogenous parameters over the whole horizon (e.g., demand, generation costs). In particular, the carbon price is assumed to coincide with the shadow price from the GEP model (see Section 3.1).

[40] These authors do not claim to be exhaustive, and neither do we.

A2. Agents have perfect information on endogenous future investments and retirements (see the anticipation module defined in Section 2.2.1).

Finally, for the fifth assumption, we need to find an alternative equivalent formulation since our model does not represent long-term and insurance markets. Here, we follow Newbery (2018) who argues – drawing on Newbery and Stiglitz (1981) – that the theoretical requirement of market completeness can be substituted by the combination of rational expectations and risk-neutrality. Because long-term markets for electricity are incomplete or 'missing' in practice (e.g., Rodilla et al., 2015; Newbery, 2016; Abada et al., 2019; Keppler et al., 2022; Wolak, 2022), this substitution also makes sense for our assessment of EOM performance. We thus make the following assumption:

A3. Agents have risk-neutral preferences.

As explained in Section 2.2, when **A1**, **A2** and **A3** hold, the representative investor behaves rationally with perfect information, fully anticipates its future decisions and understands the interplay with its current decisions, and implements the optimum in the spirit of a rational expectations equilibrium. Before relaxing **A1**, **A2** or **A3**, we turn to the issue of capacity indivisibility.

4.1.2 EOM outcomes with capacity indivisibility

To explore the impacts of discrete vs. infinitesimal capacity unit sizes, we run sensitivities with respect to the unit step size around the reference of 500 MW used in the main analysis (Section 4.2). Specifically, we assume that **A1**–**A3** hold and consider step sizes of 250, 750 and 1,000 MW. These can be interpreted as the typical sizes of investment projects, or as the volumes of disclosed projects that induce market participants to update their long-term price anticipation.

Figure 8 depicts the simulated capacity trajectories. Intuitively, the smaller is the capacity step size, the closer the simulated trajectories for an EOM are to the optimal GEP solution on average. As the step size increases, the deviation from the optimum is characterized by a delay in Solar and Storage investments and in the fossil fleet phaseout. Indicators given on average in Table 2 or on an annual basis in Figure 9 further reveal that a larger step size is conducive to larger total system cost, wholesale prices, emission levels and loss of load.[41]

Indivisibility issues in the electricity industry have long been recognized and analyzed in the literature, already in the seminal contributions by Boiteux (1949, 1960) and Williamson (1966) as well as in ensuing discussions (e.g., Andersson and Bohman, 1985). Yet they address these issues in a regulated utility environment and mainly focus on practical pricing policy considerations. More recently, Keppler (2017) reviews the literature and discusses the implications of indivisibility in a market context, observing that it leads to under-investment when coupled with inelastic demand.

[41] The LOLE increases but remains within acceptable bounds as per regulations in liberalized electricity markets (usually between 2 and 4 hours per year in expectation). Yet the increase in the LOLE explains part of the increase in average wholesale prices: as a rule of thumb, one hour of VoLL pricing at 15,000 $/MWh is tantamount to an increase in the annual average baseload price in the order of 2 $/MWh (see e.g. Figure 9).

Our results add to this literature by illustrating how indivisibility also hampers the joint dynamic of fossil phaseout with investment in renewable and storage units to achieve decarbonization.

In fact, this reveals a coordination (or circularity) issue between new entrants and existing (and possibly exiting) assets. Recall how the SD model's decision module proceeds, gradually selecting the most profitable available asset entry or exit up to an end state characterized by a zero-profit condition (Section 2.2.3). Incentives thus decrease in size as the iterative process progresses, and coordination issues start to materialize as the end state nears. For instance, in the neighborhood of optimal capacity, a potential new entry without a simultaneous exit typically leads to over-capacity that deters the actual entry decision. This is notably true for Storage that has a large contribution during scarcity hours if Peakers are not retired in a coordinated way. In turn, this further affects Solar that has to be associated with Storage to mitigate the revenue cannibalization effect.

Importantly, the materiality of this effect depends on the sensitivity of the entry and exit signals near the optimum rather than on the capacity step size relative to the overall size of the system. Appendix D illustrates that the long-term profit of a marginal investment in a given technology can be interpreted as the total cost function's gradient component with respect to installed capacity for this technology. Because these gradients are steep and asymmetric in the neighborhood of optimal capacity, this effect remains tangible even in relatively large systems.[42] As is the case here, this effect translates into extreme price sensitivity around optimal capacity, with prices jumping from a few hundreds to a few thousands $ per MWh in some peak hours. This echoes previous results in the literature highlighting that the EOM is intrinsically prone to "*erratic*" (Cramton and Stoft, 2005) price movements or "*discontinuity*" (Kraan et al., 2019).

It is worth making one final observation regarding computation time, which increases steeply as capacity step size decreases. Specifically, in our setup, reducing step size from 1,000 (resp. 500) to 500 (resp. 100) MW increases computation time by a factor of 3 (resp. 5). In a loose sense, this illustrates how long-run efficiency in an EOM is informationally and computationally demanding, especially as capacity step size gets smaller to approach GEP-style optimization with continuous capacity adjustment. That is, one realizes the amount of computations needed and information that must be processed, assessed and updated when unpacking the anticipation framework inherent to an EOM in order to attain long-run efficiency.

[42] To capture this effect, Anderson and Zachary (2023) use an approximation with costs increasing exponentially (resp. linearly) for (resp. over–) under-capacity around optimal capacity, in line with an approach implemented by the British authorities since 2016 (National Grid, 2022).

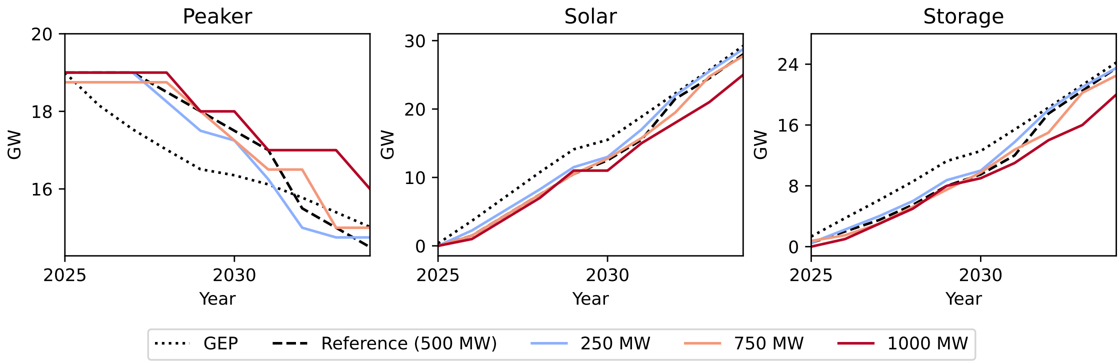

Figure 8: Capacity trajectories with different capacity step sizes (EOM with **A1**, **A2** & **A3**)

Note: Although capacity trajectories are not perfectly ordered by capacity step size (they intersect for some years), the ordering holds on average (in terms of distance to the GEP trajectory).

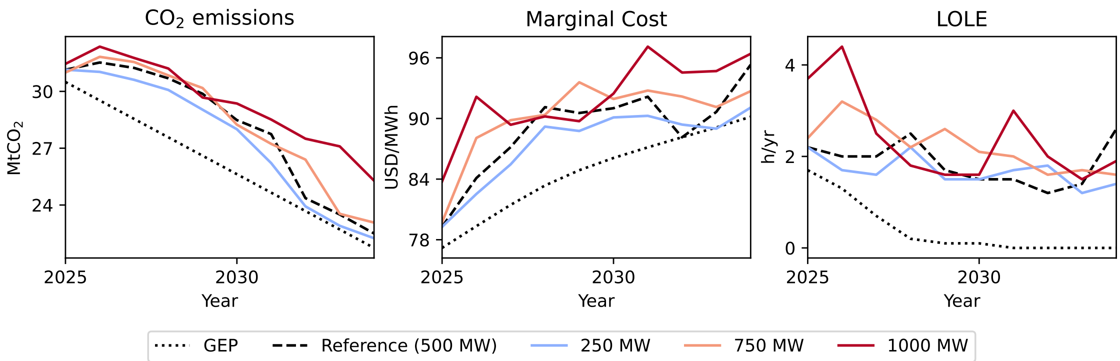

Figure 9: Annual indicators with different capacity step sizes (EOM with **A1**, **A2** & **A3**)

		Capacity step size			
	GEP	250 MW	500 MW (**ref**)	750 MW	1,000 MW
Annual Total Cost [10^9 USD/yr]	8.71	8.74	8.75	8.76	8.77
Marginal Cost [USD/MWh]	84.5	87.7	89.1	90.4	92.2
Annual Emissions [$MtCO_2$/yr]	26.1	27.5	28.1	28.4	29.4
LOLE [h/yr]	0.39	1.67	1.85	2.20	2.37
CRR Peaker [%]	32	36	56	43	48
CRR CCGT [%]	91	97	108	105	115
CRR PV [%]	100	101	105	101	104
CRR Storage [%]	100	102	108	105	112

Table 2: Average indicators with different capacity step sizes (EOM with **A1**, **A2** & **A3**)

4.2 EOM outcomes with relaxed assumptions

Let us now relax assumptions in turn and separately, keeping the capacity step size constant at 500 MW, and compare simulated EOM outcomes with those in our reference case where **A1**–**A3** jointly hold. For brevity, the case where we drop **A1** (Case 1) is relegated to Appendix F. Indeed, the impacts of downward biased anticipations of future carbon prices in Case 1 are qualitatively similar to those when we consider risk aversion and drop **A3** below (Case 3). Table 3 contains the definition of the different cases and the associated assumptions.

	Reference (Sec. 4.1.2)	Case 1 (App. F)	Case 2 (Sec. 4.2.1)	Case 3 (Sec. 4.2.2)
A1: Perfect information on exogenous parameters	✓	✗	✓	✓
A2: Perfect information on endogenous future investments and retirements	✓	✓	✗	✓
A3: Risk-neutral preferences	✓	✓	✓	✗

Table 3: Definition of studied cases and associated assumptions

4.2.1 Case 2: Biased anticipation of future entries and exits (only A2 does not hold)

In Case 2, we drop **A2** while retaining **A1** and **A3**. That is, we consider that investors and asset owners make incorrect anticipations about future investment and retirement decisions. Specifically, we consider three cases. In the first case, labeled 'static', there is no anticipation of future entries and exits (see Section 2.2.1). In the other two cases, future entries and exits are anticipated but in contrast to the perfect reference case, anticipations deviate from the optimal capacity trajectories: in the 'overestimate' (resp. 'underestimate') case, we skew positively (resp. negatively) the optimal future dynamics of Solar and Storage development, all else being equal. For illustration, we choose a + (resp. −) 40% factor bias on annual capacity additions.[43]

The 'static' and 'understimate' cases exhibit faster development for Solar and Storage early on relative to the perfect anticipation case (Figure 10). This is because the anticipation of no or lower future capacity additions magnifies expected future market revenues and in turn the incentive to invest today. That is, future capacity additions that will have a dampening effect on the whole price distribution and number of scarcity hours prices are not accounted for in full when assessing future earnings of assets possibly being invested into today. In turn, cost recovery is lower than initially expected and than in the perfect case for all assets (Table 4). This situation also illustrates a practical coordination issue across investors that is inherent to an EOM (e.g., herding behavior when investment conditions are good overall, difficulty to anticipate others' investment decisions) and possibly conducive to boom-and-bust cycles (e.g., Arango and Larsen, 2011; Hill, 2021).[44]

Over-investment is particularly salient for the first investment vintages (i.e., in the first years of the period) while the deviation is then gradually reduced over time with capacities being close to their optimal levels in the final year (Figure 10). Note that this convergence results from a modeling edge effect due to the model's finite horizon (see Section 2.2.3): as the end of the simulation horizon nears, ever less future decisions have to be anticipated, which mechanically reduces the impact of the anticipation bias. Hence, our results can only be meaningfully interpreted in the beginning of the simulation period. On the face of it, this faster early development may be seen as a positive

[43] Notice that the 'static' case can be seen as an extreme version of the 'underestimate' case. Moreover, the size of the bias only affects the quantitative nature of our results.

[44] Despite initial over-investment, a cycle does not emerge in our case study, notably because demand is structurally growing over time. This strongly mitigates potential under-profitability of assets normally linked to over-investment, and reinforces a natural asymmetry between investment and retirement decisions – the former is evaluated against both fixed OPEX and CAPEX whereas the latter only against fixed OPEX (i.e., economic conditions must turn out to be asymmetrically worse than expected to justify decommissioning).

outcome as emissions and prices are lower than in the reference case (Figure 12). However, because the first investment vintages turn out not to recoup cost in full due to biased anticipation (Figure 11), this situation will risk being economically unsustainable further down the road.

Symmetrically, the 'overestimate' case exhibits slower development for Solar and Storage early on relative to the perfect anticipation case (Figure 10). Anticipating inflated capacity additions in the future reduces expected future prices and weakens investment incentives today, hence the under-investment. But because actual capacity development happens to be lower than anticipated, cost recovery ratios are well above 100% for new assets (Figure 11). At the same time, under-investment threatens security of supply (the LOLE reaches close to 10 h/yr) and drives up prices (Table 4). Under-investment is also not conducive to reducing emissions, jeopardizing decarbonization targets.

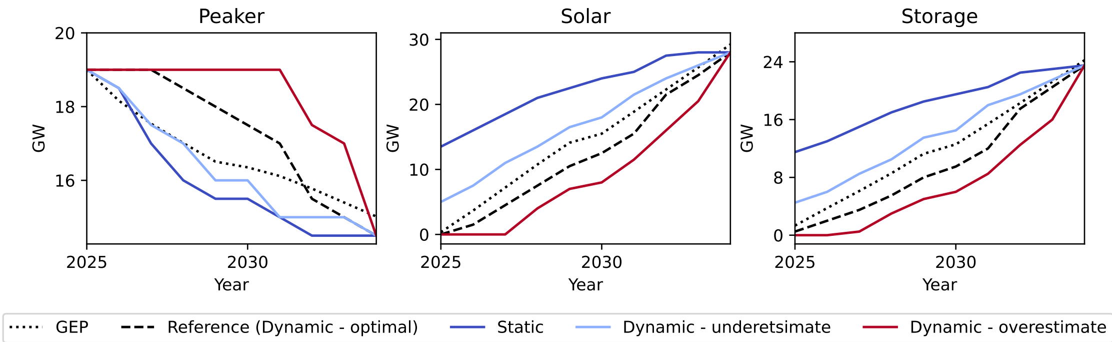

Figure 10: Capacity trajectories with different entry-exit anticipations (EOM with **A1** & **A3**)

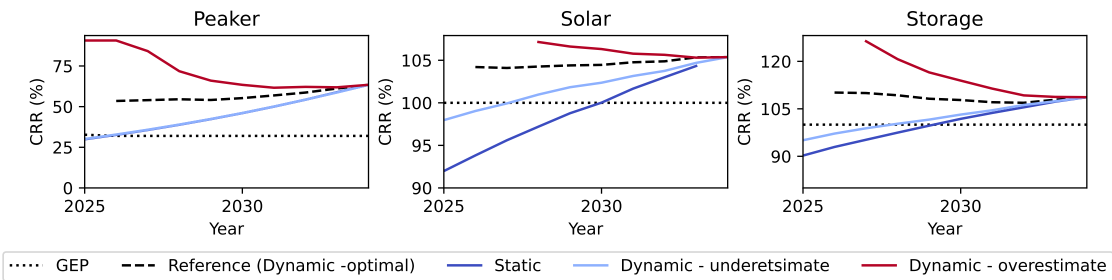

Figure 11: Cost recovery by vintage with different entry-exit anticipations (EOM with **A1** & **A3**)

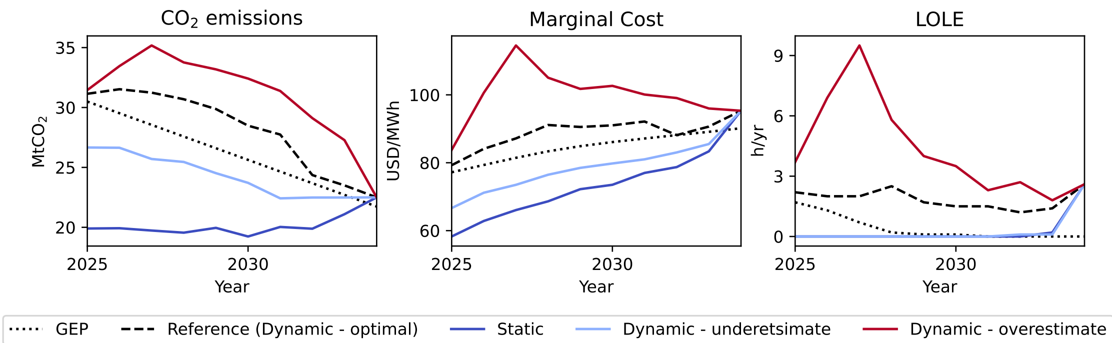

Figure 12: Annual indicators with different entry-exit anticipations (EOM with **A1** & **A3**)

	GEP	Future entry-exit anticipation: perfect (**ref**)	static	underestim.	overestim.
Annual Total Cost [10^9 USD/yr]	8.71	8.75	8.90	8.75	8.87
Marginal Cost [USD/MWh]	84.5	89.1	74.1	79.4	99.9
Annual Emissions [$MtCO_2$/yr]	26.1	28.1	20.2	24.3	31.0
LOLE [h/yr]	0.39	1.85	0.30	0.30	4.21
CRR Peak [%]	32	56	44	44	72
CRR CCGT [%]	91	108	92	97	120
CRR PV [%]	100	105	100	103	106
CRR Storage [%]	100	108	102	104	111

Table 4: Average indicators with different entry-exit anticipations (EOM with **A1** & **A3**)

4.2.2 Case 3: Risk aversion (only A3 does not hold)

In Case 3, we drop **A3** while retaining **A1** and **A2**. That is, we consider that investors and asset owners are averse to risk about future revenues and apply the certainty-equivalent decision-making criterion presented in Section 2.2.3.[45] Specifically, we consider that future market revenues $\mathbf{r}$ are uniformly distributed between 0 and $2\bar{r}$.[46] Because we lack empirical guidance to discipline the selection of a relevant risk aversion coefficient α, we follow Petitet et al. (2017) and consider a range of integer values for α between 0 and 3, where $\alpha = 0$ coincides with risk neutrality whereas $\alpha = 3$ is deemed to capture a high degree of risk aversion.[47] As we will see, the value of α affects our quantitative results monotonically, hence not their qualitative nature.

Capacity trajectories in Figure 13 exhibit a pattern of under-investment in Solar and Storage in conjunction with a delay in fossil phaseout. Risk aversion also has a noticeable impact on total system cost, ranging from +3% ($\alpha = 1$) to +14% ($\alpha = 3$). Closely looking into the central case $\alpha = 2$ shows that the delay in investment and retirement yields net savings on fixed costs (both fixed OPEX and CAPEX) of −18%, but increased fossil generation hikes total variable cost by 20% and the cost of rationing rises by +4%, totaling a net total cost increase by circa 7%.

Table 5 shows other important insights for an EOM under risk aversion. First, decarbonization is not achieved, with emissions at the end of the simulation horizon possibly even higher than at the beginning. This is because demand, which is increasing over time because of electrification in our case study, remains in an inadequate part served by fossil units instead of low-carbon units. Second, security of supply deteriorates, with both under-investment and the LOLE increasing with the degree of risk aversion (LOLE $\geq$ 10 h/yr for $\alpha \geq 2$). Third, price spikes (including hours of VoLL pricing) entail significant financial transfers between consumers and producers. Higher prices for consumers also result from risk premiums that producers need to secure to run operations.

[45] Recall that our approach to modeling risk aversion is in part grounded on the absence of sufficient, adequate hedging instruments (see Section 4.1.1).

[46] Appendix E shows that our results are qualitatively unaltered when we change the variance and the type of the probability distribution. One may also conjecture that our results would be amplified if we considered stronger forms of uncertainty and corresponding preference representation theorems. Deep uncertainty indeed prevails in the long run, especially in the context of decarbonization where the long-term energy mix, market conditions and price distributions remain largely elusive for now (e.g., Keppler et al., 2022).

[47] Following an applied study by RTE (2018) for France, we consider $\alpha = 2$ as a central value.

The effect of risk aversion has been studied in the literature with numerical models, essentially with a focus on security of supply and under-investment, and in turn on the scope and design of capacity remuneration mechanisms (e.g., Ousman Abani et al., 2018; Petitet et al., 2017; Fraunholz et al., 2023). Our results support previous findings and add to the literature on decarbonization aspects. Specifically, under-investment resulting from risk aversion hinders the existing fossil fleet phaseout which, especially when combined with growing demand for electricity, leads to emission levels jeopardizing if not even undershooting on decarbonization targets.[48]

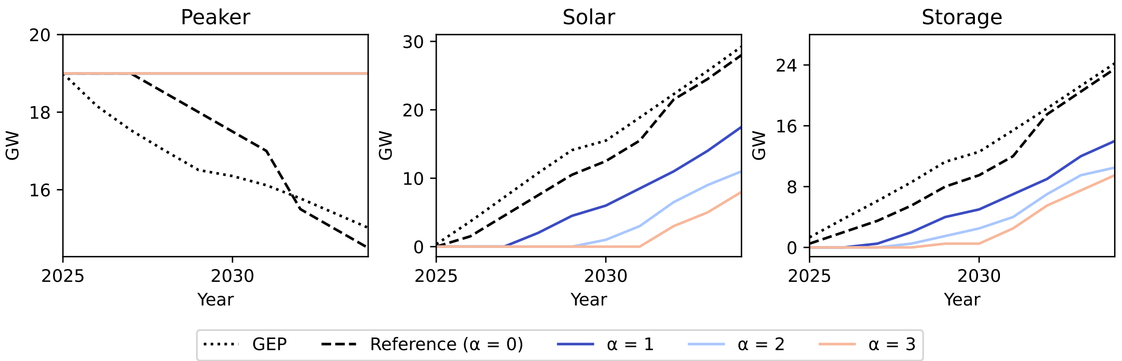

Figure 13: Capacity trajectories with different risk aversion degrees (EOM with **A1** & **A2**)

		Risk aversion degree			
	GEP	$\alpha = 0$ (**ref**)	$\alpha = 1$	$\alpha = 2$	$\alpha = 3$
Annual Total Cost [10^9 USD/yr]	8.71	8.75	8.98	9.33	9.61
Marginal Cost [USD/MWh]	84.5	89.1	108	125	136
Annual Emissions [$MtCO_2$/yr]	26.1	28.1	33.5	36.1	37.4
LOLE [h/yr]	0.39	1.85	5.67	10.0	13.4
CRR Peak [%]	32	56	80	214	301
CRR CCGT [%]	91	108	163	243	290
CRR PV [%]	100	105	120	146	170
CRR Storage [%]	100	108	129	162	182

Table 5: Average indicators with different risk aversion degrees (EOM with **A1** & **A2**)

4.3 Summary of modeling results and policy implications

Practical limits of price-based coordination. Several conclusions emerge from our modeling approach and results. First, the implementation and unpacking of the EOM framework clarify the assumptions needed to reproduce the optimal long-run outcome from a GEP model. By contrast, these theoretical assumptions and their practical implications are often implicit or remain elusive in the literature. Notably, we highlight the high level of informational and computational complexity associated with optimal anticipations of all relevant future market fundamentals and future entry and exit decisions. Assuming optimal anticipations, rational decision-making and risk neutrality,

[48] One may also conjecture that capturing the feedback loop between price and demand dynamics in the long run (not represented in our framework) would exacerbate the shortcomings of the EOM for decarbonization. For instance, high or volatile electricity prices could be detrimental to electrification, either directly by deterring investment in electrical equipment or indirectly by limiting public support if low-carbon generation investment does not keep pace.

	GEP	Simulated case: Reference *(indivisibility)*	Case 2 *(biased anticipation)*	Case 3 *(risk aversion)*
Annual Total Cost [10^9 USD/yr]	8.71	8.74–8.77	8.75–8.90	8.98–9.61
Marginal Cost [USD/MWh]	84.5	87.7–92.2	74.1–99.9	108–137
Annual Emissions [$MtCO_2$/yr]	26.1	27.5–29.4	20.2–31.0	33.5–37.4
LOLE [h/yr]	0.39	1.67–2.37	0.30–4.21	5.67–13.4
CRR Peaker [%]	32	36–48	44–72	80–301
CRR CCGT [%]	91	97–115	92–120	163–290
CRR PV [%]	100	101–104	100–106	120–170
CRR Storage [%]	100	102–112	102–111	129–182

Table 6: Summary of average indicators across simulations (range of values is provided)

we further illustrate how the convergence of the EOM outcomes towards the optimum occurs only when the capacity step size is small enough – ideally infinitely small –, highlighting the issue of lumpy investment and retirement decisions. While the total system cost is not very sensitive to unit indivisibility, this causes a coordination issue between new low-carbon investments and the phaseout of fossil assets that affects carbon emissions more markedly.

In the second part of our analysis, we relax these assumptions in isolation and find significant deviations (see Table 6 for a summary of quantitative indicators). Intuitively, the direction of the deviation in investment depends on the direction of the bias in the anticipation of future market entries – an upward bias relative to the optimum leads to a downward bias in anticipated market revenues that delays investment decisions, and vice versa. Moreover, on top of the standard result of under-investment when investors are risk averse, our joint modeling of entry and exit decisions also shows how decarbonization targets risk not being met if the existing fossil fleet is not pushed out of the market economically by new low-carbon entrants in due time.

Potential benefits of complementary quantity-based coordination. Our modeling framework constitutes a good basis for extensions to explore changes in market design that could improve long-run coordination and risk-sharing mechanisms (e.g., Keppler et al., 2022). Different add-ons such as a long-term contracting module could be plugged into our core model, of which a variety of designs could be assessed and compared. To illustrate, let us consider the case where assumptions **A1–A3** hold and the optimal fossil phaseout is exogenously enforced, although we do not specify at this stage through which mechanism it is implemented. A cursory examination of the results in Figure 14 shows that capacity trajectories are closer to the optimum. Importantly, this holds not just for Peakers whose phaseout is exogenously driven, but also for Solar and Storage. This shows how explicit coordination through quantities has the potential to complement implicit coordination through prices and improve on market efficiency in the long run.

Of course, the extent to which this potential can be tapped into depends on the implementation mechanism. One may for instance think of a situation where the regulator steers the decommission-

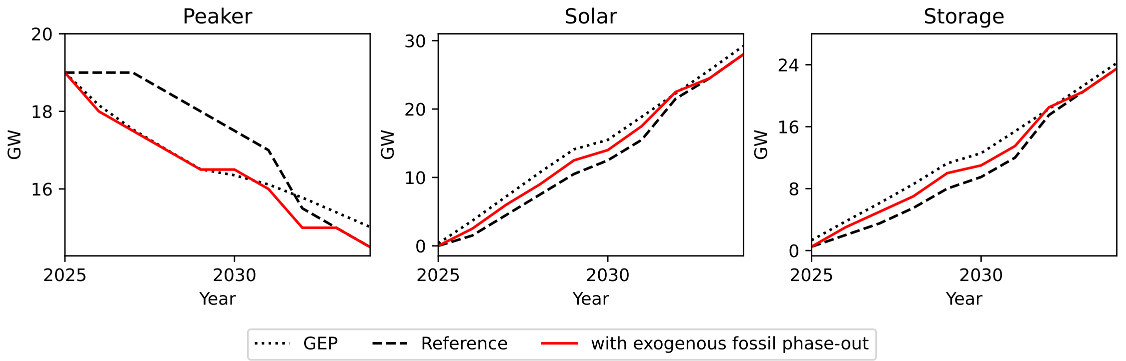

Figure 14: Capacity trajectories with vs. without optimal phaseout (EOM with **A1**, **A2** & **A3**)

ing trajectory through an auction scheme, similarly to what is done in Germany (e.g., Tiedemann and Müller-Hansen, 2023). But various alternatives could also be investigated, notably regarding the form of compensation payments awarded in the auction or the type of planning used to define the phaseout path enforced by the regulator (in a bid to avoid the pitfall of substituting an 'imperfect' market mechanism by 'perfect' regulation). Similar crucial design options also exist in the case of long-term contracting mechanisms for investment, be they government-backed contracts issued through public auctions or private contracts stimulated through the provision of public guarantees or through an obligation on retailers (see notably CEPR, 2023, for a review).

5 Conclusion

In these times of renewed debates on electricity market design in the context of decarbonization, the EOM has often been criticized for various, at times dubious reasons. For instance, the increasing share of generation with (near) zero short-term marginal cost has been blamed for reducing prices and (expected) asset profitability. However, this merit-order effect is either transitory (Antweiler and Muesgens, 2021) or caused by inadequate policy choices (Brown and Reichenberg, 2021) rather than reflective of a limitation inherent to the EOM. In fact, price distributions are more likely to change in shape than to be lowered on average (e.g., Ekholm and Virasjoki, 2020; Mallapragada et al., 2021). Rather, our results suggest that a key issue with the EOM paradigm is that long-run efficiency holds only under a set of idealistic assumptions. When one of these preconditions is not met, entry-exit coordination through market price signals alone is insufficient and deviations from the optimum occur (see Section 4.3 for a summary). This notably leads to higher electricity prices, lower security of supply, and higher emission levels that imperil decarbonization.

Different market design reform options are being debated at the time of writing, primarily in Europe. Our modeling framework can be extended in different ways to inform these debates and ensuing policy choices. For instance, assessing and comparing design alternatives for a long-term contracting module (as delineated e.g. in Joskow, 2022; Keppler et al., 2022; Wolak, 2022) would be particularly relevant. This includes various approaches to auction design (e.g., Iossa et al., 2022;

Fabra and Montero, 2023), contract design (e.g., Billimoria and Simshauser, 2023; Newbery, 2023) and planning (e.g., Corneli, 2020; Anderson and Zachary, 2023). In this respect, accounting for realistic behavioral, informational and structural assumptions will be of the essence.

Data and code availability

The Python source code and setup instructions are provided here: GitHub/ANTIGONE.
Datasets and a detailed documentation are provided here: Zenodo/Dataset.

Acknowledgements

We are indebted to Yannick Perez, Olivier Massol, Marie Girod, Vincent Rious, Adrien Nicolle and Jean-Yves Bourmaud for their comments and suggestions. We would also like to thank participants at the 1st IAEE Online Conference (2021), the 2021 TransAtlantic Infraday, the 2021 FAEE Student Workshop and the 13th TSE Conference on The Economics of Energy and Climate (2022) for useful remarks. The usual disclaimer applies.
Funding from Électricité de France is gratefully acknowledged. The views expressed in this article are solely those of the authors and do not necessarily reflect those of EDF.

References

Abada, I., de Maere d'Aertrycke, G., Ehrenmann, A., Smeers, Y., 2019. What models tell us about long-term contracts in times of the energy transition. Economics of Energy & Environmental Policy 8, 163–181.

Abada, I., de Maere d'Aertrycke, G., Smeers, Y., 2017. On the multiplicity of solutions in generation capacity investment models with incomplete markets: A risk-averse stochastic equilibrium approach. Mathematical Programming 165, 5–69.

Abdin, A., Caunhye, A., Zio, E., Cardin, M.A., 2022. Optimizing generation expansion planning with operational uncertainty: A multistage adaptive robust approach. Applied Energy 306, 118032.

Acharya, V., Lochstoer, L.A., Ramadorai, T., 2013. Limits to arbitrage and hedging: Evidence from commodity markets. Journal of Financial Economics 109, 441–465.

Ahmad, S., Tahar, R.M., Muhammad-Sukki, F., Munir, A.B., Rahim, R.A., 2016. Application of system dynamics approach in electricity sector modelling: A review. Renewable and Sustainable Energy Reviews 56, 29–37.

Alimou, Y., Maïzi, N., Bourmaud, J.Y., Li, M., 2020. Assessing the security of electricity supply through multi-scale modeling: The TIMES-ANTARES linking approach. Applied Energy 279, 115717.

Anderson, E., Zachary, S., 2023. Minimax decision rules for planning under uncertainty: Drawbacks and remedies. European Journal of Operational Research 311, 789–800.

Andersson, R., Bohman, M., 1985. Short– and long-run marginal cost pricing: On their alleged equivalence. Energy Economics 7, 279–288.

Antweiler, W., Muesgens, F., 2021. On the long-term merit order effect of renewable energies. Energy Economics 99, 105275.

Anwar, M.B., Stephen, G., Dalvi, S., Frew, B., Ericson, S., Brown, M., O'Malley, M., 2022. Modeling investment decisions from heterogeneous firms under imperfect information and risk in wholesale electricity markets. Applied Energy 306, 117908.

Arango, S., Larsen, E., 2011. Cycles in deregulated electricity markets: Empirical evidence from two decades. Energy Policy 39, 2457–2466.

Arrow, K.J., Debreu, G., 1954. Existence of an equilibrium for a competitive economy. Econometrica 22, 265–290.

Baumol, W.J., Quandt, R.E., 1964. Rules of thumb and optimally imperfect decisions. The American Economic Review 54, 23–46.

Bessembinder, H., Lemmon, M.L., 2002. Equilibrium pricing and optimal hedging in electricity forward markets. The Journal of Finance 57, 1347–1382.

Billimoria, F., Simshauser, P., 2023. Contract design for storage in hybrid electricity markets. EPRG Working Paper 2304. University of Cambridge.

Blaquez, J., Fuentes, R., Manzano, B., 2020. On some economic principles of the energy transition. Energy Policy 147, 111807.

Boiteux, M., 1949. La tarification des demandes en pointe : Application de la théorie de la vente au coût marginal. Revue Générale de l'Électricité 58, 321–340.

Boiteux, M., 1960. Peak-load pricing. The Journal of Business 33, 157–179.

Brock, W.A., Hommes, C.H., 1997. A rational route to randomness. Econometrica 65, 1059–1095.

Brown, T., Reichenberg, L., 2021. Decreasing market value of variable renewables can be avoided by policy action. Energy Economics 100, 105354.

Bruninx, K., Ovaere, M., Delarue, E., 2020. The long-term impact of the market stability reserve on the EU emission trading system. Energy Economics 89, 104746.

Bublitz, A., Keles, D., Zimmermann, F., Fraunholz, C., Fichtner, W., 2019. A survey on electricity market design: Insights from theory and real-world implementations of capacity remuneration mechanisms. Energy Economics 80, 1059–1078.

Bunn, D., Dyner, I., 1996. Systems simulation to support integrated energy analysis and liberalised planning. International Transactions in Operational Research 3, 105–115.

Burtraw, D., Holt, C., Palmer, K., Shobe, W., 2022. Price-responsive allowance supply in emissions markets. Journal of the Association of Environmental and Resource Economists 9, 851–884.

CAISO, 2018. Production and curtailments historical data. http://www.caiso.com/Documents/ProductionAndCurtailmentsData_2018.xlsx. California Independent System Operator.

CAISO, 2019. Production and curtailments historical data. http://www.caiso.com/Documents/ProductionAndCurtailmentsData_2019.xlsx. California Independent System Operator.

CAISO, 2022. 2021 annual report on market issues and performance. California Independent System Operator.

California State Senate, 2018. The 100 Percent Clean Energy Act of 2018. https://www.energy.ca.gov/sb100. Senate Bill 100.

Caramanis, M., 1982. Investment decisions and long-term planning under electricity spot pricing. IEEE Transactions on Power Apparatus and Systems 101, 4640–4648.

Cason, T.N., Stranlund, J.K., de Vries, F.P., 2023. Investment incentives in tradable emissions markets with price floors. Journal of the Association of Environmental and Resource Economists 10, 283–314.

Cepeda, M., Finon, D., 2011. Generation capacity adequacy in interdependent electricity markets. Energy Policy 39, 3128–3143.

CEPR, 2023. Electricity market design: Views from European economists. Ambec, S., Banal, A., Cantillon, E., Crampes, C., Creti, A., Decarolis, F., Fabra, N., Gerlagh, R., Neuhoff, K., Landais, C., Liski, M., Llobet, G., Newbery, D., Polo, M., Reguant, M., Schwenen, S., Vehviläinen, L. Policy Insight 120. Centre for Economic Policy Reseach.

Chen, H., Wang, C., Cai, W., Wang, J., 2018. Simulating the impact of investment preference on low-carbon transition in power sector. Applied Energy 217, 440–455.

Cicala, S., 2022. Imperfect markets versus imperfect regulation in US electricity generation. The American Economic Review 112, 409–441.

Corneli, S., 2020. A Prism based configuration market for rapid low cost and reliable electricity sector decarbonization. Working Paper. World Resources Institute.

CPUC, 2019. RESOLVE model and results used for 2019 IRP portfolio development. https://www.cpuc.ca.gov/General.aspx?id=6442462824. California Public Utilities Commission.

Cramton, P., Stoft, S., 2005. A capacity market that makes sense. The Electricity Journal 18, 43–54.

Creti, A., Fontini, F., 2019. Economics of Electricity – Markets, Competition and Rules. Cambridge University Press.

Dixit, A.K., Pindyck, R.S., 1994. Investment under Uncertainty. Princeton University Press.

Dyner, I., Ochoa, C., Franco, C.J., 2011. SD for assessing electricity market integration. Proceedings. The 29th International Conference of the System Dynamics Society.

Ehrenmann, A., Smeers, Y., 2011. Generation capacity expansion in a risky environment: A stochastic equilibrium analysis. Operations Research 59, 1332–1346.

Ekholm, T., Virasjoki, V., 2020. Pricing and competition with 100% variable renewable energy and storage. The Energy Journal 42, 1–17.

European Commission, 2019. The European Green Deal. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM:2019:640:FIN. COM(2019) 640 final.

European Commission, 2022. Fit for 55 Package. https://www.consilium.europa.eu/en/policies/green-deal/fit-for-55-the-eu-plan-for-a-green-transition.

Fabra, N., 2023. Reforming European electricity markets. Working Paper. Centre for Economic Policy Research.

Fabra, N., Montero, J.P., 2023. Technology-neutral vs. technology-specific procurement. The Economic Journal 133, 669–705.

Fan, L., Norman, C.S., Patt, A.G., 2012. Electricity capacity investment under risk aversion: A case study of coal, gas, and concentrated solar power. Energy Economics 34, 54–61.

Felder, F., 2002. The need for governance of restructured electric power systems and some policy implications. The Electricity Journal 15, 36–43.

Ford, A., 1983. Using simulation for policy evaluation in the electric utility industry. SIMULATION 40, 85–92.

Fraunholz, C., Kraft, E., Keles, D., Fichtner, W., 2021. Advanced price forecasting in agent-based electricity market simulation. Applied Energy 290, 116688.

Fraunholz, C., Miskiw, K.K., Kraft, E., Fichtner, W., Weber, C., 2023. On the role of risk aversion and market design in capacity expansion planning. The Energy Journal 44, 111–138.

Friedrich, M., Mauer, E.M., Pahle, M., Tietjen, O., 2020. From fundamentals to financial assets: The evolution of understanding price formation in the EU ETS. Working Paper. Leibniz Information Centre for Economics.

Froot, K.A., Scharfstein, D.S., Stein, J.C., 1993. Risk management: Coordinating corporate investment and financing policies. The Journal of Finance 48, 1629–1658.

Fuss, S., Flachsland, C., Koch, N., Kornek, U., Knopf, B., Edenhofer, O., 2018. A framework for assessing the performance of cap-and-trade systems: Insights from the European Union emissions trading system. Review of Environmental Economics and Policy 12, 220–241.

Gabaix, X., 2014. A sparsity-based model of bounded rationality. The Quarterly Journal of Economics 129, 1661–1710.

GitHub/ANTIGONE, 2023. ANTIGONE source code GitHub repository. https://github.com/alexis-lebeau-research/ANTIGONE.

Gohdes, N., Simshauser, P., Wilson, C., 2022. Renewable entry costs, project finance and the role of revenue quality in Australia's National Electricity Market. Energy Economics 114, 106312.

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

Iossa, E., Rey, P., Waterson, M., 2022. Organising competition for the market. Journal of the European Economic Association 20, 822–868.

Jagannathan, R., Matsa, D.A., Meier, I., Tarhan, V., 2016. Why do firms use high discount rates? Journal of Financial Economics 120, 445–463.

Joskow, P.L., 2008. Lessons learned from electricity market liberalization. The Energy Journal 29, 9–42.

Joskow, P.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, 313–329.

Joskow, P.L., Léautier, T.O., 2021. Optimal wholesale pricing and investment in generation: The basics. Edward Elgar Publishing, Cheltenham, UK. chapter 3. pp. 36–72.

Joskow, P.L., Schmalensee, R., 1988. Markets for power: An analysis of electrical utility deregulation. MIT Press.

Joskow, P.L., Tirole, J., 2007. Reliability and competitive electricity markets. The RAND Journal of Economics 38, 60–84.

Kagiannas, A.G., Askounis, D.T., Psarras, J., 2004. Power generation planning: A survey from monopoly to competition. International Journal of Electrical Power & Energy Systems 26, 413–421.

Keles, D., Bublitz, A., Zimmermann, F., Genoese, M., Fichtner, W., 2016. Analysis of design options for the electricity market: The German case. Applied Energy 183, 884–901.

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

Keppler, J.H., Quemin, S., Saguan, M., 2022. Why the sustainable provision of low-carbon electricity needs hybrid markets. Energy Policy 171, 113273.

Kraan, O., Kramer, G.J., Nikolic, I., Chappin, E., Koning, V., 2019. Why fully liberalised electricity markets will fail to meet deep decarbonisation targets even with strong carbon pricing. Energy Policy 131, 99–110.

Littlechild, S.C., 1988. Spot pricing of electricity: Arguments and prospects. Energy Policy 16, 398–403.

Lonergan, K.E., Egli, F., Osorio, S., Sansavini, G., Pahle, M., Schmidt, T.S., , Steffen, B., 2023. Improving the representation of cost of capital in energy system models. Joule 7, 1–15.

Léautier, T.O., 2019. Imperfect markets and imperfect regulation. MIT press.

Mallapragada, D.S., Junge, C., Wang, C.X., Pfeifenberger, J., Joskow, P.L., Schmalensee, R., 2021. Electricity Price Distributions in Future Renewables-Dominant Power Grids and Policy Implications. Working Paper 29510. National Bureau of Economic Research.

Mays, J., Jenkins, J., 2022. Electricity markets under deep decarbonization. Working Paper No. 22-550. United States Association for Energy Economics.

Mays, J., Morton, D.P., O'Neill, R.P., 2019. Asymmetric risk and fuel neutrality in electricity capacity markets. Nature Energy 4, 948–956.

Montgomery, W.D., 1972. Markets in licenses and efficient pollution control programs. Journal of Economic Theory 5, 395–418.

National Grid, 2022. Electricity capacity report. National Grid ESO.

Neuhoff, K., May, N., Richstein, J.C., 2022. Financing renewables in the age of falling technology costs. Resource and Energy Economics 70, 101330.

Neuhoff, K., de Vries, L.J., 2004. Insufficient incentives for investment in electricity generations. Utilities Policy 12, 253–267.

Newbery, D.M., 2016. Missing money and missing markets: Reliability, capacity auctions and interconnectors. Energy Policy 94, 401–410.

Newbery, D.M., 2018. What future(s) for liberalized electricity markets: Efficient, equitable or innovative? The Energy Journal 39, 1–28.

Newbery, D.M., 2023. Efficient renewable electricity support: Designing an incentive-compatible support scheme. The Energy Journal 44, 1–22.

Newbery, D.M., Stiglitz, J.E., 1981. The theory of commodity price stabilization: A study in the economics of risk. Oxford University Press.

Ninja Renewables, 2021. Database. https://www.renewables.ninja/.

Osorio, S., Tietjen, O., Pahle, M., Pietzcker, R., Edenhofer, O., 2021. Reviewing the market stability reserve in light of more ambitious EU ETS emission targets. Energy Policy 158, 112530.

Ousman Abani, A., 2019. Electricity market design for long-term capacity adequacy in a context of energy transition. Ph.D. thesis. PSL Research University.

Ousman Abani, A., Hary, N., Rious, V., Saguan, M., 2018. The impact of investors' risk aversion on the performances of capacity remuneration mechanisms. Energy Policy 112, 84–97.

Peluchon, B., 2021. Decarbonizing the electricity sector efficiently requires a change of market design. CEEM Working Paper 51. Paris-Dauphine University.

Perino, G., Willner, M., Quemin, S., Pahle, M., 2022. The European Union emissions trading system market stability reserve: Does it stabilize or destabilize the market? Review of Environmental Economics and Policy 16, 338–345.

Petitet, M., Janssen, T., Finon, D., 2016. Carbon price instead of support schemes: Wind power investments by the electricity market. The Energy Journal 37, 109–140.

Petitet, M., Janssen, T., Finon, D., 2017. Capacity adequacy in power markets facing energy transition: A comparison of scarcity pricing and capacity mechanism. Energy Policy 103, 30–46.

Pollitt, M., 2021. The future design of the electricity market, in: Handbook on Electricity Markets. Edward Elgar Publishing. chapter 16, pp. 428–442.

Poncelet, K., Kaminski, S., Delarue, E., 2019. Ability and limitations of optimization models for computing the equilibrium in competitive energy markets. TME Working Paper EN019-09. KU Leuven.

Pourramezan, A., Samadi, M., 2023. A system dynamics investigation on the long-term impacts of demand response in generation investment planning incorporating renewables. Renewable and Sustainable Energy Reviews 171, 113003.

Quemin, S., Pahle, M., 2023. Financials threaten to undermine the functioning of emissions markets. Nature Climate Change 13, 22–31.

Quemin, S., Trotignon, R., 2021. Emissions trading with rolling horizons. Journal of Economic Dynamics and Control 125, 104099.

Reis, R., 2006. Inattentive producers. The Review of Economic Studies 73, 793–821.

Rios-Festner, D., Blanco, G., Olsina, F., 2019. Integrating real options analysis with long-term electricity market models. Energy Economics 80, 188–205.

Rios-Festner, D., Blanco, G., Olsina, F., 2020. Long-term assessment of power capacity incentives by modeling generation investment dynamics under irreversibility and uncertainty. Energy Policy 137, 111185.

Rodilla, P., Batlle, C., 2012. Security of electricity supply at the generation level: Problem analysis. Energy Policy 40, 177–185.

Rodilla, P., García-González, J., Álvaro Baíllo, Cerisola, S., Batlle, C., 2015. Hydro resource management, risk aversion and equilibrium in an incomplete electricity market setting. Energy Economics 51, 365–382.

Roques, F., Finon, D., 2017. Adapting electricity markets to decarbonisation and security of supply objectives: Toward a hybrid regime? Energy Policy 105, 584–596.

RTE, 2018. Analyse d'impact du mécanisme de capacité. Réseau de Transport d'Électricité.

RTE, 2021. Energy pathways to 2050: Key results. Réseau de Transport d'Électricité.

Rubin, J.D., 1996. A model of intertemporal emission trading, banking, and borrowing. Journal of Environmental Economics and Management 31, 269–286.

Ruhnau, O., Bucksteeg, M., Ritter, Schmitz, R., Böttger, D., Koch, M., Pöstges, A., Wiedmann, M., Hirth, L., 2022. Why electricity market models yield different results: Carbon pricing in a model-comparison experiment. Renewable and Sustainable Energy Reviews 153, 111701.

dos Santos, A.F., Saraiva, J.T., 2021. Agent based models in power systems: A literature review. U. Porto Journal of Engineering 7, 101–113.

Schennach, S.M., 2000. The economics of pollution permit banking in the context of Title IV of the 1990 Clean Air Act amendments. Journal of Environmental Economics and Management 40, 189–210.

Schittekatte, T., Batlle, C., 2023. Power crisis in the EU 3.0: Proposals to complete long-term markets. Working Paper 2023-04. MIT Centre for Energy and Environmental Policy Research.

Schulthoff, M., Rudnick, I., Bose, A., Gençer, E., 2021. Role of hydrogen in a low-carbon electric power system: A case study. Frontiers in Energy Research 8, 585461.

Schweppe, F.C., Caramanis, M.C., Tabors, R.D., Bohn, R.E., 1988. Spot pricing of electricity. Kluwer Academic Publishers.

Simon, H.A., 1955. A behavioral model of rational choice. The Quarterly Journal of Economics 69, 99–118.

Tang, L., Guo, J., Zhao, B., Wang, X., Shao, C., Wang, Y., 2021. Power generation mix evolution based on rolling horizon optimal approach: A system dynamics analysis. Energy 224, 120075.

Tao, Z., Moncada, J.A., Delarue, E., 2023. Exploring the impact of boundedly rational power plant investment decision-making by applying prospect theory. Utilities Policy 82, 101522.

Tao, Z., Moncada, J.A., Poncelet, K., Delarue, E., 2021. Review and analysis of investment decision making algorithms in long-term agent-based electric power system simulation models. Renewable and Sustainable Energy Reviews 136, 110405.

Teufel, F., Miller, M., Genoese, M., Fichtner, W., 2013. Review of system dynamics models for electricity market simulations. Working Paper. Leibniz Information Centre for Economics.

Tiedemann, S., Müller-Hansen, F., 2023. Auctions to phase out coal power: Lessons learned from Germany. Energy Policy 174, 113387.

Tvinnereim, E., Mehling, M., 2018. Carbon pricing and deep decarbonisation. Energy Policy 121, 185–189.

Van Huyck, J.B., Battalio, R.C., Beil, R.O., 1990. Tacit coordination games, strategic uncertainty, and coordination failure. The American Economic Review 80, 234–248.

Ventosa, M., Baıllo, A., Ramos, A., Rivier, M., 2005. Electricity market modeling trends. Energy Policy 33, 897–913.

de Vries, L.J., Chappin, E., Richstein, J., 2013. EMLab-Generation – An experimentation environment for electricity policy analysis. Working Paper. Delft University of Technology.

de Vries, L.J., Heijnen, P., 2008. The impact of electricity market design upon investment under uncertainty: The effectiveness of capacity mechanisms. Utilities Policy 16, 215–227.

Vázquez, C., Rivier, M., Perez-Arriaga, I.J., 2002. A market approach to long-term security of supply. IEEE Transactions on Power Systems 17, 349–357.

Weber, C., Heidari, S., Bucksteeg, M., 2021. Coping with uncertainties in the electricity sector – Methods for decisions of different scope. Economics of Energy & Environmental Policy 10, 5–30.

Westfield, F.M., 1988. Electric power: Competition or centralization? Energy Policy 16, 378–383.

Willems, B., Morbee, J., 2010. Market completeness: How options affect hedging and investments in the electricity sector. Energy Economics 32, 786–795.

Williamson, O.E., 1966. Peak-load pricing and optimal capacity under indivisibility constraints. The American Economic Review 56, 810–827.

Williamson, O.E., 1975. Markets and hierarchies: Analysis and antitrust implications. New York: Free Press.

Wolak, F.A., 2022. Long-term resource adequacy in wholesale electricity markets with significant intermittent renewables. Environmental and Energy Policy and the Economy 3, 155–220.

Zenodo/Dataset, 2023. ANTIGONE California dataset. https://zenodo.org/record/8177393.

Zimmermann, F., Keles, D., 2023. State or market: Investments in new nuclear power plants in France and their domestic and cross-border effects. Energy Policy 173, 113403.

Appendices

A Notations and units

Sets and indices	
$\mathcal{H}$	Set of hours in a year, indexed by h
$\mathcal{Y}$	Set of years, indexed by y
$\mathcal{W}$	Set of weather scenarios, indexed by w
$\mathcal{G}$	Set of conventional dispatch technologies, indexed by g
$\mathcal{V}$	Set of variable renewable energy technologies, indexed by v
$\mathcal{S}$	Set of storage technologies, indexed by s
$\mathcal{T}$	Set of all technologies ($\mathcal{T} = \mathcal{G} \cup \mathcal{V} \cup \mathcal{S}$), indexed by t
$\mathcal{U}_t$	Set of units of technology t, indexed by u
Parameters and variables	
β	Discount factor
Δ	Time step duration (here, one hour)
Π_w	Probability of weather scenario w
$D_{y,w,h}$	Load in year y, weather scenario w and hour h [MW]
$\lambda_{y,w,h}$	Marginal cost of electricity in year y, weather scenario w and hour h [USD/MWh]
$OC_{y,t}$	Annual fixed O&M cost for technology t [$/MW/yr]
$IC_{y,t}$	Investment cost annuity for technology t [$/MW/yr]
$VC_{y,t}$	Annual operating cost for technology t [$/MWh]
γ_t	Carbon intensity of technology t [tCO_2/MWh]
Q_y	Annual CO_2 emissions cap [tCO_2]
ℓ_t	Lifespan of technology t [yr]
$\mathcal{L}_u$	Year of initially scheduled closure of unit u
$\mathcal{R}_{y,u}$	Net market revenues in year y for unit u
$n_{y,t}$	Number of operating units in year y for technology t
$n^+_{y,t}$	Number of developed units in year y for technology t
$n^-_{y,t}$	Number of closed units in year y for technology t
$\kappa_{y,t}$	Total installed capacity in year y for technology t [MW]
$\alpha_{t,w,h}$	Hourly availability of technology t [%]
k_t	Power capacity of technology t [MW/unit]
$q_{t,y,w,h}$	Production of technology t in year y, weather scenario w and hour h [MW]
$c_{s,y,w,h}$	Power charged into technology s in year y, weather scenario w and hour h [MW]
$soc_{s,y,w,h}$	State of charge of technology s in year y, weather scenario w and hour h [MWh]
ρ_s	Charging and discharging efficiency of technology s [%]
d_s	Storage duration for technology s [hours]
$f_{y,w,h}$	Lost load in year y, weather scenario w and hour h [MW]
VoLL	Value of Lost Load [$/MWh]

Table A.1: Model notations and units

B GEP model description

This Appendix describes and interprets the equations and constraints of the generation expansion planning (GEP) optimization model using notations given in Appendix A.

The objective function is the expected discounted total cost over the planning horizon, that is

$$\min_{\{n,n^+,n^-,q,f,c\}} \sum_{y\in\mathcal{Y}} \beta^y \left\{ \sum_{w\in\mathcal{W}} \Pi_w \sum_{h\in\mathcal{H}} \left[\sum_{t\in\mathcal{T}} VC_{y,t}\cdot q_{t,y,w,h} + \text{VoLL}\cdot f_{y,w,h} \right] + \sum_{t\in\mathcal{T}} \left[OC_{y,t}\cdot n_{y,t} + IC_{y,t}\cdot n^+_{y,t}\cdot \sum_{i=0}^{\min(\ell_t,\#\mathcal{Y}-y)} \beta^i \right] \right\}$$

where $\#\mathcal{X}$ denotes the cardinality of set $\mathcal{X}$. This formulation accommodates: conventional dispatchable generation units characterized by variable generation costs and availability profiles; variable renewable units with zero variable cost and hourly capacity factors; short-term storage units with power and energy components linked by duration and round-trip efficiency parameters. Storage units are modeled deterministically and dispatched across time steps assuming intertemporal arbitrage with perfect foresight. Each technology is represented by discrete homogeneous units (i.e., the decision variables are expressed in terms of number of units).

The first set of constraints (B.1–B.4) represent the hourly dispatch, that is $\forall y\in\mathcal{Y}, w\in\mathcal{W}$,

$$\forall h\in\mathcal{H},\ \sum_{t\in\mathcal{T}} q_{t,y,w,h} + f_{y,w,h} = D_{y,w,h} + \sum_{s\in\mathcal{S}} c_{s,y,w,h}, \tag{B.1}$$

$$\forall h\in\mathcal{H}, t\in\mathcal{T},\ q_{t,y,w,h} \le k_t \alpha_{t,w,h} n_{y,t}, \tag{B.2}$$

$$\forall h\in\mathcal{H}, s\in\mathcal{S},\ soc_{s,y,w,h} \le k_s d_s n_{y,s}, \tag{B.3}$$

$$\forall h\in\mathcal{H}^\star, s\in\mathcal{S},\ soc_{s,y,w,h} = soc_{s,y,w,h-1} + \rho_s c_{s,y,w,h-1} - q_{s,y,w,h-1}/\rho_s, \tag{B.4}$$

where (B.1) imposes load balance, (B.2) imposes the upper limit on generation (for simplicity dynamic generation constraints such as ramp-up rates are not represented), (B.3) imposes the upper limit on stored energy, and (B.4) reflects the storage dynamics with round-trip efficiency.

The second set of constraints (B.5–B.6) represent the fleet dynamics, that is

$$\forall y\in\mathcal{Y}^\star, h\in\mathcal{H}, t\in\mathcal{T},\ n_{y,t} = n_{y-1,t} + n^+_{y,t} - n^-_{y,t}, \tag{B.5}$$

$$\forall y\in\mathcal{Y}, t\in\mathcal{T},\ \text{if } y+\ell_t \le \#\mathcal{Y}:\ \sum_{i=y}^{\#\mathcal{Y}} n^-_{i,t} \ge n^+_{y,t}, \tag{B.6}$$

where (B.5) tracks the number of units per technology over time and (B.6) imposes that each endogenous investment can be associated with a decommissioning decision during its lifespan.

Third, constraint (B.7) imposes an annual cap on CO_2 emissions whose trajectory $\{Q_y\}_y$ is exogenously given, that is

$$\forall y \in \mathcal{Y}, \ \sum_{w \in \mathcal{W}} \Pi_w \sum_{t \in \mathcal{T}} \sum_{h \in \mathcal{H}} \gamma_t \cdot q_{t,y,h} \leq Q_y. \tag{B.7}$$

Each decision variables can be constrained in an *ad-hoc* manner with an upper/lower bound or with a specific value. This feature is used to model the existing fleet for which n can be fixed at the beginning of the simulation (the initial fleet described in Section 3) and n^+ can be constrained to 0 afterwards if the technology is not available for new developments.

Finally, all decision variables (i.e., n, n^+, n^-, q, f, c) have non-negativity constraints.

C Decision algorithm description

The structure of the investment and decommissioning decision algorithm for the representative agent in our System Dynamics market simulator is sketched below.

Algorithm 1 Decision module in the SD market simulator

for $y \in \mathcal{Y}$ **do**
  Remove all units reaching the end of their lifespan
  Initialize empty array U_{decom} to store decommissioned units during decision loop
  Initialize empty array U_{invest} to store invested units during decision loop
  Initialize empty array S to store successive states of the fleet
  Form the set of anticipations for representative agents (see Section 2.2.1)
  continue = True

  while continue **do** ▷ *decision loop*
    Create an empty array D to store NPV of all possible decisions

    for $t \in \mathcal{T}$ **do** ▷ *assess new investment options*
      if t is eligible for investment **then**
        Compute NPV of a new project over its whole lifespan, including full CAPEX
        if $NPV > 0$ **then**
          Store NPV value associated with this investment decision in D

    for $t \in \mathcal{T}$ **do** ▷ *assess decommissioning decision for units in existing fleet*
      if t is eligible for decommissioning **then**
        Compute net revenues $\mathcal{R}_{y,t}$ in year y (infra-marginal rent minus fixed OPEX)
        if $\mathcal{R}_{y,t} < 0$ **then**
          Compute NPV over the remaining lifespan, considering avoidable costs (i.e., fixed OPEX)
          if $NPV < 0$ **then**
            Store NPV value associated with this decomissioning decision in D

    for $u \in U_{\text{decom}}$ **do** ▷ *assess postponing closures decided in previous iterations*
      Compute NPV of running the asset for one extra year, considering avoidable costs (i.e., fixed OPEX and not CAPEX)
      if $NPV > 0$ **then**
        Store NPV value associated with this closure postponement in D

    for $u \in U_{\text{invest}}$ **do** ▷ *assess renouncing to investments decided in previous iterations*
      Compute net revenues $\mathcal{R}_{y,u}$ in year y (infra-marginal rent minus fixed OPEX and annualized CAPEX)
      if $\mathcal{R}_{y,u} < 0$ **then**
        Compute NPV over the remaining lifespan, including full CAPEX
        if $NPV < 0$ **then**
          Store NPV value associated with this investment renouncement in D

    if D is not empty **then** ▷ *stopping criterion*
      Pick and implement decision with highest NPV in absolute terms
      if new state of the fleet is already in S **then**
        continue = False
      else
        Store the state of the fleet in S
    else
      continue = False

D Convergence of the simulation model

This Appendix lays out some theoretical considerations regarding the convergence of the simulation model. For clarity and without loss of generality, we consider a simplified setup with one representative year, one weather scenario, continuous capacity adjustments, full availability, no storage and no fixed O&M costs. This simplified case helps build intuition, and its logic and resolution extend to more complex cases. If we consider the representative year as a steady state, our optimization problem can be interpreted as a long-term cost-minimization problem that writes

$$\min_{\{\kappa_t\},\{q_{t,h}\},\{f_h\}} \mathcal{C}(\kappa_t, q_{t,h}, f_h) = \sum_{t\in\mathcal{T}} IC_t \cdot \kappa_t + \sum_{h\in\mathcal{H}} \left[\text{VoLL} \cdot f_h + \sum_{t\in\mathcal{T}} VC_t \cdot q_{t,h} \right]$$

subject to

$$f_h + \sum_{t\in\mathcal{T}} q_{t,h} = \Delta \cdot D_h \quad (\lambda_h), \text{ and}$$

$$q_{t,h} \le \Delta \cdot \kappa_t \quad (\mu_{t,h}), \quad q_{t,h} \ge 0 \quad (\nu_{i,h}), \quad f_h \ge 0 \quad (\xi_h), \quad \kappa_t \ge 0 \quad (\pi_t),$$

where the variables within parentheses denotes the constraints' dual variables.

For any given positive values of $\{\kappa_t\}_t > 0$ (i.e., not necessarily optimal), the gradient component with respect to κ_t of the Lagrangian function (of all primal and dual variables) is given by

$$\frac{\partial \mathcal{L}}{\partial \kappa_t} = IC_t - \Delta \cdot \sum_{h\in\mathcal{H}} \mu_{t,h}.$$

If we minimize $\mathcal{C}$ for these given $\{\kappa_t\}_t$ (i.e., we solve the dispatch problem for a given capacity vector), we can define the set of hours $\mathcal{H}_t^+$ where technology t is infra-marginal and we get from the dispatch problem's KKT conditions

$$\sum_{h\in\mathcal{H}} \tilde{\mu}_{t,h} = \sum_{h\in\mathcal{H}_t^+} (\tilde{\lambda}_h - VC_t).$$

Combining the above equations, it comes

$$\frac{\partial \mathcal{L}}{\partial \kappa_t} = IC_t - \Delta \cdot \sum_{h\in\mathcal{H}_t^+} (\tilde{\lambda}_h - VC_t) = -LTP,$$

where LTP is the long-term profit. We thus see that the normalized long-term profit calculated for a given technology in an out-of-equilibrium state (in terms of installed capacity) corresponds to the Lagrangian function's gradient component with respect to the capacity installed for this technology. Therefore, the iterative procedure defining the simulation model in Appendix C can mathematically be interpreted as a *coordinate gradient descent* algorithm in an ideal case. While it may not be the most computationally-efficient approach to solve this linear problem, it does have the merit to be meaningful from an economic point of view.

E Certainty-equivalent formulation

This Appendix provides details on the approach and calculations of the certainty equivalent used in the decision module under risk aversion and uncertain aggregate market revenues $\mathbf{r}$ (Section 2.2.3). By definition, the certainty equivalent $r^\star$ is the certain revenue that yields the same utility as the expected utility over the random distribution of revenues, i.e. $\mathcal{U}(r^\star) \equiv \mathbb{E}\{\mathcal{U}(\mathbf{r})\}$. With the functional form for $\mathcal{U}$ defined in Section 2.2.3, we can compute $\mathcal{U}(r^\star)$ and infer $r^\star$ assuming specific probability density functions f_R for $\mathbf{r}$. The overall approach is sketched in Figure E.1.

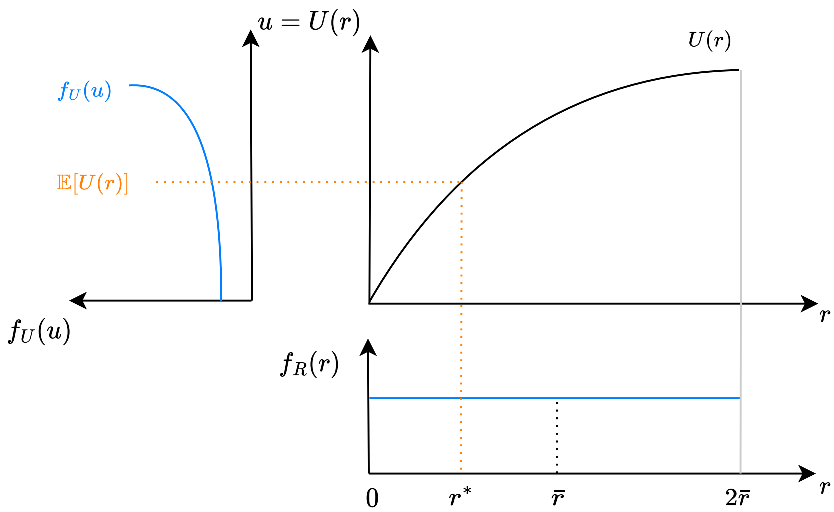

Figure E.1: Certainty equivalent calculation when revenues $\mathbf{r}$ are uniformly distributed over $[0, 2\bar{r}]$

We consider two cases as in Neuhoff et al. (2022). For the central case of a uniform distribution with finite support $[0, 2\bar{r}]$ used in Section 4.2, we get

$$\mathcal{U}(r^\star) = \int_{-\infty}^{+\infty} \mathcal{U}(\mathbf{r}) f_R(\mathbf{r}) \mathrm{d}\mathbf{r} = 1 + \frac{\exp(-2\alpha) - 1}{2\alpha} \Rightarrow r^\star = -\frac{\bar{r}}{\alpha} \cdot \ln\left(\frac{1 - \exp(-2\alpha)}{2\alpha}\right).$$

Alternatively, we consider a normal distribution with finite mean μ and variance σ^2. In this case, we have that $\mathcal{U}(r^\star) = 1 - \mathbb{E}\{\exp(-\alpha \mathbf{r}/\bar{r})\}$, where the second term on the right-hand side is the mean of a log-normal distribution. We thus have

$$\mathcal{U}(r^\star) = 1 - \exp\left(-\alpha + \frac{1}{2}\left(\frac{\alpha\sigma}{\bar{r}}\right)^2\right) \Rightarrow r^\star = \bar{r} - \frac{\alpha}{2\bar{r}}\sigma^2.$$

Next, we explore numerically how the certainty equivalent varies with the degree of risk aversion and the variance for both types of distribution. Figure E.2 shows that the certainty equivalent to expected revenue ratio $r^\star/\bar{r}$ decreases with α, and that it is always larger in the case of a uniform distribution. This implies that considering a normal distribution would amplify the results with a uniform distribution in Section 4.2, especially for high values of α.

Additionally, a higher variance in the sense of a mean-preserving spread is conducive to a lower

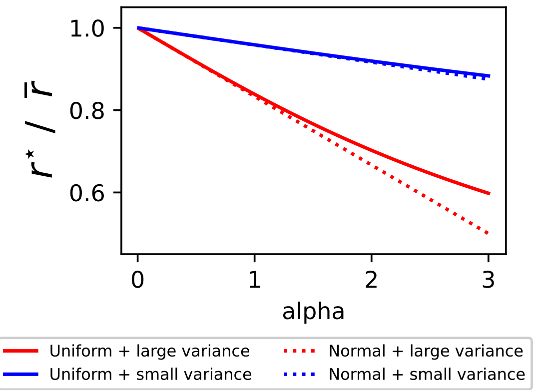

Figure E.2: Certainty equivalent to expected revenue ratio under different modeling assumptions

$r^\star/\bar{r}$ ratio. Specifically, we consider the case where the variance varies by a factor of four. For the uniform distribution, this is tantamount to reducing the support by a factor of two: the variance is $\bar{r}^2/3$ with the $[0, 2\bar{r}]$ support vs. $\bar{r}^2/12$ with the $[\bar{r}/2, 3\bar{r}/2]$ support. For the normal distribution, we simply adjust the variance parameter accordingly. With this calibration, Figure E.2 illustrates that the qualitative nature of the results in Section 4.2 is unaltered by the type and variance of the probability distribution for aggregate market revenues.

F Biased carbon price anticipation (Case 1)

This Appendix provides additional simulations for the case of downward biased anticipations of future carbon prices. Case 1 is appended because the results are qualitatively similar to those with risk aversion (Case 3, Section 4.2.2), at least in the first years of the simulations before the 'edge effect' materializes. It also relies on a more arbitrary, less micro-founded modeling choice.

In Case 1, we drop **A1** while retaining **A2** and **A3**. That is, we consider that investors and asset owners make conservative carbon price forecasts relative to the optimal trajectory satisfying annual emissions targets. Motivations for this assumption are threefold. First, prices in existing carbon markets have by and large been too low or volatile to convey robust long-term investment signals in line with these targets (e.g., Tvinnereim and Mehling, 2018; Perino et al., 2022). Second, market imperfections or regulatory distortions including limited foresight, excessive discounting or insufficient policy credibility may distort price formation and anticipation downwards in the short to mid term (e.g., Fuss et al., 2018; Quemin and Trotignon, 2021). Third, carbon price formation may be driven by various factors other than fundamentals, making it difficult to predict future prices (e.g., Friedrich et al., 2020; Quemin and Pahle, 2023).

We consider that for each year in the simulation, the current price coincides with the optimal one from the GEP model, but that the representative agent anticipates that the price will grow at a lower rate than in the optimal trajectory. Specifically, we consider three cases for the anticipated annual growth rate of the carbon price (CAGR), namely 0, 2 and 4% compared to the reference case with an optimal growth rate of around 6%. Figure F.1 shows that the lower the CAGR is, the more capacity entries and exits are delayed. Recall that we found similar results by increasing the degree of risk aversion (see Case 3, Section 4.2.2), but the driving mechanism is different. This time, the delay originates from a biased anticipation of competitive advantage (resp. disadvantage) tilted towards (resp. against) fossil-fired plants (resp. solar and storage assets).

Additionally, note that by the same token as in Case 2 (Section 4.2.1), all the capacity trajectories converge towards the reference case with unbiased carbon price anticipation at the end of the horizon due to a factitious edge effect (i.e., anticipations are by construction less and less biased the nearer the end of the simulation period due to the modeling artifact whereby the last year of the simulation period is repeated until assets' lifetimes are covered in whole). To illustrate this further, whereas emissions are equal across all CAGRs on the last year of the simulation period (because the installed asset fleet and market conditions are the same), the delay induced by a lower CAGR results in higher emissions over the whole period (Table F.1).

Finally, the cost recovery analysis reveals extra revenues for all asset types, which are increasing with the anticipation bias (i.e., decreasing with the CAGR). This is because the realized carbon price is actually higher than anticipated, which increases the realized price of electricity on average and is overall economically beneficial across the whole asset fleet. Intuitively, Figure F.2 shows that this effect is more pronounced early on in simulation period (i.e., when a given anticipation bias has a greater impact on entry and exit decisions, all else being equal).

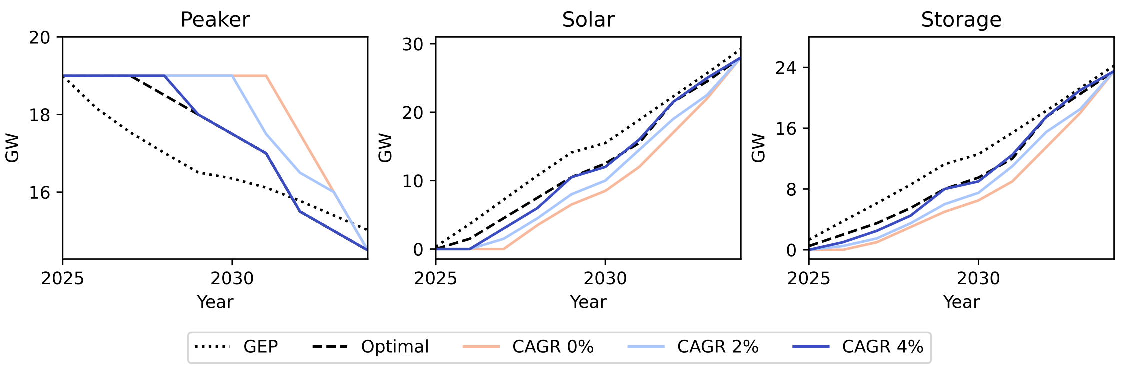

Figure F.1: Capacity trajectories with different anticipation biases (EOM with **A2** & **A3**)

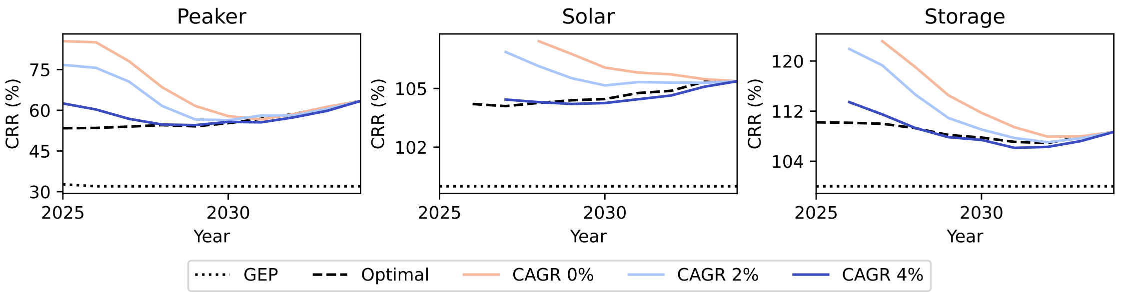

Figure F.2: Cost recovery by vintage with different anticipation biases (EOM with **A2** & **A3**)

		CAGR			
	GEP	0 %	2 %	4 %	6 % (**ref**)
Annual Total Cost [10^9 USD/yr]	8.71	8.84	8.80	8.76	8.74
Marginal Cost [USD/MWh]	84.5	98.3	95.2	90.9	87.9
Annual Emissions [$MtCO_2$/yr]	26.1	30.6	29.7	28.5	27.1
LOLE [h/yr]	0.39	3.79	3.22	2.33	2.02
CRR Peaker [%]	32	68	64	58	56
CRR CCGT [%]	91	118	114	109	106
CRR PV [%]	100	106	105	105	104
CRR Storage [%]	100	110	109	108	106

Table F.1: Average indicators with different anticipation biases (EOM with **A2** & **A3**)