Risk Implications of the Deployment of Renewables for Investments in Electricity Generation

by

Fernando J. de Sisternes

Ingeniero Industrial, Universidad Carlos III de Madrid (2005)
S.M. Technology and Policy, Massachusetts Institute of Technology (2010)

Submitted to the Engineering Systems Division
in partial fulfillment of the requirements for the degree of
DOCTOR OF PHILOSOPHY
at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
June 2014

Author .....
Engineering Systems Division
May 27, 2014

Certified by .....
Mort D. Webster
Assistant Professor of Engineering Systems
Thesis Supervisor

Certified by .....
Ignacio J. Pérez-Arriaga
Visiting Professor of Engineering Systems
Thesis Supervisor

Certified by .....
Richard Schmalensee
Howard W. Johnson Professor of Management, Emeritus
Professor of Economics, Emeritus
Dean Emeritus
Committee Member

Certified by .....
John E. Parsons
Senior Lecturer, Sloan School of Management
Head, MBA Finance Track
Executive Director, Center for Energy and Environmental Policy Research
Committee Member

Accepted by .....
Richard C. Larson
Professor, Engineering Systems Division
Chair, Engineering Systems Division Education Committee

To my parents

Luis Fernando and María del Carmen

Risk Implications of the Deployment of Renewables for Investments in Electricity Generation

by

Fernando J. de Sisternes

Submitted to the Engineering Systems Division
on May 27, 2014, in partial fulfillment of the
requirements for the degree of
DOCTOR OF PHILOSOPHY

Abstract

This thesis explores the potential risk implications that a large penetration of intermittent renewable electricity generation –such as wind and solar power– may have on the future electricity generation technology mix, focusing on the anticipated new operating conditions of different thermal generating technologies and their remuneration in a competitive market environment. In addition, this thesis illustrates with an example how risk should be valued at the power plant level in order to internalize the potential risks to which the generators are exposed.

This thesis first compares the impacts of three different bidding rules on wholesale prices and on the remuneration of units in power systems with a significant share of renewable generation. The effects of bidding rules are distinguished from the effects of regulatory uncertainty that can unexpectedly increase renewable generation by considering two distinct situations: 1) an 'adapted' capacity mix, which is optimized for any given amount of renewable penetration, and 2) a 'non-adapted' capacity mix, which is optimized for zero renewable penetration, but that operates with a certain non-zero renewable capacity, added on top of an already adequate system. The analysis performed stresses the importance of sound mechanisms that allow the full-cost recovery of plants in a system where the intermittency of renewables accentuates nonconvex costs, without over-increasing the cost paid by consumers for electricity. Additionally, the analysis quantifies the potential losses incurred by different thermal technologies if renewable deployment occurs without allowing for adaptation.

Methodologically, this thesis uses a novel long-term generation investment model, the Investment Model for Renewable Electricity Systems (IMRES), to determine the minimum cost thermal capacity mix necessary to complement renewable generation to meet electricity demand, and to extract hourly wholesale prices. IMRES is a capacity expansion model with unit commitment constraints whose main characteristics are: 1) reflecting the impact of hourly resolution operation constraints on investment decisions and on total generation cost; 2) accounting for the chronological variability of demand and renewable output, and the correlation between the two; and 3) deciding on power plant investments at the individual

plant level. These characteristics allow for a detailed analysis of the profits obtained by individual plants in systems with large renewable penetration levels. In addition, this thesis tests the performance of a heuristic method that selects four weeks from a full year series to optimally represent the net load duration curve (i.e., the difference between demand and renewable output, decreasingly ordered). For each application of this heuristic method, three metrics are proposed to reflect that the approximation also represents the chronological variability of the net load.

Lastly, this thesis explores the role of risk in the valuation of electricity generating technologies and shows how to incorporate standard risk pricing principles into the popular Monte Carlo simulation analysis. The exposition is structured using the standard framework for a typical Monte Carlo cash flow simulation so that the implementation can be readily generalized. This framework stresses the necessity of an asset pricing approach to assess the relationship between the risk in the assets cash flows and the macroeconomic risk with which the financial investors are concerned. The framework provided is flexible and can accommodate many different structures for the interaction between the macroeconomic risk and the risks in the asset's cash flows (such as those from shocks in renewable deployment).

Thesis Supervisor: Mort D. Webster
Title: Assistant Professor of Engineering Systems

Thesis Supervisor: Ignacio J. Pérez-Arriaga
Title: Visiting Professor of Engineering Systems

Committee Member: Richard Schmalensee
Title: Howard W. Johnson Professor of Management, Emeritus
Professor of Economics, Emeritus
Dean Emeritus

Committee Member: John E. Parsons
Title: Senior Lecturer, Sloan School of Management
Head, MBA Finance Track
Executive Director, Center for Energy and Environmental Policy Research

Acknowledgments

This thesis would have never been possible without the support from many people.

First and foremost, I would like to thank my thesis supervisors, Ignacio Pérez-Arriaga and Mort Webster; and my other doctoral committee members, John Parsons and Richard Schmalensee, for their guidance during these four years and for finding time in their busy schedules to meet with me. Special thanks to Ignacio for generously sharing his infinite knowledge on power systems regulation, for being a great mentor, for his friendship, his candor, and for the productive late night meetings in the cafes of Cambridge. Special thanks to Mort for being a permanent source of invaluable advice, an amazing mentor and for giving me the opportunity to work together teaching power systems models. Thanks to John for teaching me everything I know about risk management, for his patience, and his friendship; and thanks to Richard for his critical insight and his invaluable feedback. It was a great privilege and an enjoyable challenge to work with this amazing team.

Special thanks to David Marks, who always had time to meet with me, even before I was admitted to MIT, and guided me into the PhD program. For that, I will always be grateful. Thanks also to Carlos Batlle for his insightful comments on how to implement the different bidding rules included in this thesis, for sharing his expert knowledge, and for his friendship. I am also very grateful to Asun González and to Ed Cook for their outpouring of kindness and for making me feel like family.

I am indebted to the institutions that have funded my graduate studies from the very beginning. Thanks to 'La Caixa' Foundation for believing in me when I had very little to show on paper and for opening the door to an amazing group individuals. Thanks also to the Caja Madrid Foundation, the Martin Family Society of Fellows for Sustainability, and the National Science Foundation for supporting me in the subsequent years of my doctorate until today.

Conceiving a model like IMRES and building its present capabilities has been a major effort that has benefitted from the aid and the comments from many people. Here I have to thank Andrés Ramos for helping me come to terms with GAMS and always being available

for questions at the inception of the model. I also thank Javier García-González for spending every year one week in the cold of the Boston winter to share his modeling expertise. The validation of IMRES and its multiple extensions occurred during my collaboration with the International Energy Agency. Thanks to Simon Müller for putting his faith in IMRES to answer the questions posed in the GIVAR project, and for motivating me to enhance the capabilities of the model; and thanks to Álvaro Portellano for encouraging me to go always one step further, for his friendship, his support and for sharing a few squash skills with me.

I am grateful for having been part of some of the extraordinary communities living at MIT, especially ESD, ESS, ESRG, CEEPR, the Webster Research Group, and Spain@MIT. Among all the people in these communities, I would like to thank especially Maite Peña, Alex Jacquillat and Berk Ustun for being amazing office mates and friends. Thanks also to Stephen Zoepf, David Ramberg, Claudia Octaviano, Vivek Sakhrani, Rebecca Saari, Bryan Palmintier, Valerie Karplus, Dan Livengood, Katherine Dykes, Nidhi Santen, Pearl Donohoo-Vallet, Jen Morris, Tommy Leung, Rhonda Jordan, Jesse Jenkins, Mar Reguant and Audun Botterud. Special thanks to José Chamorro for his friendship; and to Beth Milnes for always being available and helpful.

I feel immensely fortunate and blessed for the truly exceptional individuals that I have found on the way, and who have accompanied me during all or part of these six years: Sebastian Pfotenhauer, Adriana Ortega, Eph Lanford, Matt Cherian, Roberto Sanchís, Paula Sánchez, Carlos Pardo, Jorge Cañizales, Ana Díaz, Maité Balda, Ada Yeste, Ruben García, María de Soria, María José Nieves, Alessandro Golkar, Ferrán Vidal, Joel Saà, Maite Peña, Noel Silva, Israel Valentín, Patricia Gómez, Sara Hernández, Helena de Puig, Miguel Villalobos, Iñaki Cornago, Paula Ciaurriz, Laura Escudero, Sergi Perdices, Elisa Castañer, David Colino, Enrique Lizarraga, Enrique Fernández, Josep Marc Mingot, Marc Sánchez, Lucinio Muñoz, Laura Díaz-Anadón, Antonio Labrador, Chema Fernández and Santiago Ibañez. Rafael Montilla, my fellow baron, deserves a special mention for sharing so many good times and laughs together and, most importantly, for his visits when I was entangled writing this thesis. Daniel Selva, my friend since the Orientation Week in Indiana, also deserves spe-

cial mention for his endless generosity, for all the good times, and for being a role model. Thanks also to the many dwellers of 165 Erie for being great roommates and friends: Avi Wolfson, Stephen Woodrow, Adrian Rodriguez, Antonio Tugores, Anna Delgado, Paul van der Boor, Jan Eide, Arthur Gueneau, Michael Chang, Paul Kishimoto, Philip Caplan and Luis Alvarez. Thank you all for making such an amazing family.

Thanks to all the old friends and family that have always supported me. Thanks to Jesús, Carlos, Elisa, Bianca, Sonia, Ana, Nacho, Abel, Francesc and Eduardo for always being there.

Last but definitely not least, I would like to thank my parents, my sister, my two dear grandmas, my aunt Purita, and my family in Marbella for their unconditional and unwavering support and love. They mean the world to me and I would have never made it without them.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Summary

In restructured markets, the asymmetry in the perception of risk by generators and consumers, the volatility of prices and the presence of price caps have been considered –rightly or not– potential driving factors for underinvestment in new generating capacity (Newell, 2012), therefore requiring some kind of regulatory intervention. More recently, in systems with a high penetration of intermittent renewables[1], the potential price depression and higher volatility driven by low variable cost renewable electricity and the reduction of thermal plants capacity factor, together with a more costly and technically demanding operating regime can potentially exacerbate the problem creating a negative profit[2] expectation for certain technologies (Perez-Arriaga and Batlle, 2012).

This thesis assesses, both theoretically and by means of detailed computer simulations, the potential impact that a large penetration of intermittent renewable electricity generation –such as wind and solar power– may have on generation investments in electricity

[1]Here intermittent refers to the non-controllable variability and the partial unpredictability of renewable resources, mostly of wind power and solar photovoltaics (Perez-Arriaga and Batlle, 2012)

[2]The generation costs to be "recovered" (or not) via market prices include all the operation costs (here only startup costs and fuel costs of electricity production have been considered), plus the capital costs (amortization and a reasonable rate of return on investment). In this thesis by "profit" it is understood any excess income to a plant from market prices that is beyond strict cost recovery.

markets. Specifically, this thesis concentrates on the impact of renewable generation on the profits obtained by different generation technologies, and on the cost paid by consumers for electricity.

First, the thesis focuses on analyzing the impact of renewable generation on the optimal generation mix, and on the capacity factor, the energy contribution, the startup regime and the profits of four standard thermal generation technologies: nuclear, coal, combined-cycle gas turbines (CCGTs) and open-cycle gas turbines (OCGTs). This analysis is performed for ten different renewable penetration scenarios under two distinct situations: 1) an 'adapted' capacity mix that is optimized for the given renewable penetration level; and 2) a 'non-adapted' capacity mix, optimized for zero renewables but operated under a non-zero renewable capacity. It also includes a sensitivity analysis of the impact of the lumpiness introduced by the size of power plants, the value of the startup cost and the value of lost load (VOLL) on the results.

Second, as renewable penetration increases, so does the cycling regime of thermal power plants and the fuel and O&M costs associated with starting up the generation units. In order to help generators recover these costs, some electricity market operators have adopted complex bidding mechanisms that account explicitly for nonconvex costs that are not reflected in the marginal cost of energy, while other markets rely solely on simple bids with revenue sufficiency conditions. Examples of such mechanisms include side payments, price uplifts, and internalizing startup costs within a simple bid. This thesis also addresses this topic, studying the impact of three different bidding mechanisms on the total cost-recovery of different technologies and on the cost paid by the consumer.

Third, given the uncertainty around renewable deployment, the role of risk in the valuation of generating technologies has gained significant attention in recent years. Many of the risks to which generation assets are exposed (e.g., electricity prices, fuel prices, new regulation, etc) introduce a series of nonlinear and time-varying relationships between them and the cash flows produced by the generator. Yet, investors in power plants commonly account for risk using the risk-adjusted discounting methodology, which assumes that exposure is

proportional to the market risk, resulting in a valuation of the power plant that is prone to errors. This thesis introduces the stochastic discounting methodology for the valuation of generation assets, as a method that accurately values risk by explicitly accounting for the correlation between the average market risk and the expected cash flows generated by the power plant.

The findings of this thesis are primarily targeted at regulators and policymakers who have to confront the complexities of designing electricity market rules in the presence of intermittent renewables. In addition, the findings in this work can inform investors in power generation assets who need to understand the actual exposure to risk of this type of assets and help them to value the asset accordingly and make better investment decisions.

1.2 Background of Wholesale Electricity Markets

Since the liberalization of electricity generation began in 1982 in Chile, many countries have decided to introduce competition in their electricity industry to achieve greater economic efficiency (Joskow and Schmalensee, 1983) (Rothwell and Gomez, 2003) (Sioshansi, 2006). This process has reduced the costs of service and other inefficiencies associated with the monopolistic behavior of the traditional vertically integrated utility (Fabrizio et al., 2007). Competition is best achieved through markets where the commodity can be freely traded. Accordingly, liberalized electric power systems have established wholesale markets for electricity generation to let market forces decide on the price and quantity of electricity that is to be produced. In this context, it is the role of the regulator to establish a robust set of market rules that ensures competition while guaranteeing the security of electricity supply (Perez-Arriaga, 2007) (Rodilla, 2010) (Rodilla and Batlle, 2012) (Stoft, 2002).

Under liberalization, market agents are left with tasks that they can perform more efficiently than the central planner. Subject to the conditions of electricity wholesale markets, operating decisions are not taken by a central operator, but by individual agents who try to maximize their profits (Joskow and Schmalensee, 1983) (Schweppe et al., 1987). These agents also decide on individual power plant investments that build up the system's electricity

generation capacity. According to economic theory, under ideal conditions a remuneration scheme based only on the short-term marginal cost of the system would allow generators recover both operation costs and capital costs –including amortization and a rate of return on investment– (Kahn, 1971) (Schweppe et al., 1987) (Perez-Arriaga, 1994) (Perez-Arriaga and Meseguer, 1997), attracting enough investors to build an adequate level of generating capacity. Nevertheless, there are practical difficulties in implementing a pricing scheme that accurately reflects the consumers' real willingness to pay for electricity and the actual costs incurred by the system. These two obstacles combined with users' and investors' risk aversion can result in underinvestment or overinvestment situations, yielding a suboptimal generating capacity mix that does not reflect the desired level of generation adequacy in the system. The resulting market failure has been observed since the foundations of electricity wholesale markets (Schweppe et al., 1987), and establishing market rules that allow revenue reconciliation –the ability to recover all costs– has remained a major challenge (Batlle and Perez-Arriaga, 2008) (Rodilla and Batlle, 2012) (Schweppe et al., 1987) (Stoft, 2002).

1.2.1 The Missing Money Problem

Historically, addressing regulatory actions that limit prices from reaching the expected high levels during times of scarcity has been the most salient issue, constituting the so-called 'missing money' problem. According to Hogan (Hogan, 2005):"The missing money problem arises when occasional market price increases are limited by administrative actions such as price caps. By preventing prices from reaching high levels during times of relative scarcity, these administrative actions reduce the payments that could be applied towards the fixed operating costs of existing generation plants and the investment costs of new plants." Regulators introduce price caps in order to protect consumers against price spikes and to reduce the ability of companies to exert market power (Stoft, 2002)(Kirschen and Strbac, 2004). Yet, the 'missing money' problem only occurs during a transitory after the regulatory change: as it will be shown later, in equilibrium, for any given price limitation being it high or low, the 'missing money' problem does not exist as the generating capacity adapts to whatever

the new scarcity price is.

1.2.2 Addressing Adequacy

Adequacy refers to "the existence of enough available capacity of generation and network, both installed and/or expected to be installed, to meet demand" (Perez-Arriaga, 2007) In equilibrium, systems with a low price cap that does not reflect the actual desired level of generation adequacy have adopted regulatory solutions based on different forms of capacity mechanisms (capacity payments, capacity markets or reliability options). However, these mechanisms are decoupled from the 'missing money' problem, as their sole original objective is to enhance the systems adequacy in equilibrium. Rodilla (Rodilla, 2010) presents a comprehensive treatment of capacity mechanisms and their impact on the different dimensions of security of supply (security, firmness, adequacy and strategic energy policy).

1.2.3 Accounting for Nonconvexities in Wholesale Markets

The theoretical derivation of optimal prices assumes a convex problem where decision variables are continuous and the feasible region defined by the constraints of the problem is also convex. However, the discrete nature of some of the decisions in electricity markets (e.g., the size of plants, the commitment state of a plant, startup decisions, etc) can produce situations in which there might not exit a set of prices that support an equilibrium solution (i.e., prices that not only reflect the marginal cost of energy, but also other nonconvex costs in the system such as those from starting-up the units) (Gribik et al., 2007). In the past, this issue has had less economic importance than the effects on profits brought in by regulatory changes and it has typically attracted less attention. However, as will be shown in subsequent sections, the cost associated with cycling increases significantly in systems with large amounts of intermittent renewables. Therefore, recovering nonconvex costs, as well as understanding the implications of pricing mechanisms on profit, becomes an increasingly important problem.

When a mismatch between revenue and cost occurs, revenue reconciliation mechanisms

have to be established to make power plants whole. Schweppe et al. contemplated the possibility of introducing modifications to wholesale prices to achieve revenue sufficiency, highlighting the importance of the tradeoff between equity (having consumers pay for the full cost of generation) and economic efficiency (that the modified price is close to the optimal marginal price).

> "A key property of marginal cost based spot prices is [that] they tend to recover both operating and capital costs. Since generation is assumed to be dispatched optimally, marginal costs exceed average operating costs. Thus, charging customers at marginal costs yields revenues which exactly match operating and capital costs. Unfortunately, in the real world, this difference will usually either over-or underrecover capital costs." (Schweppe et al., 1987)

Vazquez (Vazquez, 2003) focused on finding a mechanism to obtain prices that provide the optimal economic signals in the short and in the long term when the problem is nonconvex. He found that the price resulting from the slope of the envelope of the cost function provides the result closest to optimality. He also found that the prices derived from a commitment problem where the binary variables are relaxed (prices derived from this problem are easy to obtain), coincide in most cases with the prices resulting from the slope of the envelope of the cost function.

O'Neill et al. proposed a method to obtain market clearing prices in power systems with nonconvexities that would yield zero profits (O'Neill et al., 2005). The method is based on solving a cost minimization problem, including integer decisions, in order to obtain a welfare maximizing solution; calculating a set of prices derived from the same problem but with integer decisions fixed; and offering bidders additional payments associated with their lumpy decisions.

Similarly, Gribic et al. analyzed the impact of introducing a complementary payment[3] that compensates plants for accepting a dispatch solution that does not maximize their

[3]Here 'complementary payment' refers to a lump sum paid for the difference between the operating costs incurred and the energy revenues obtained, similar to the 'side payment' bidding rule that will be discussed in following sections

profits (as a result of ignoring nonconvex costs in calculating the market-clearing price and quantities), under three different alternative conventions to calculate the market-clearing price (Gribik et al., 2007).

1.2.4 Accounting for the Uncertainty of Renewable Deployment

In addition to issues pertaining to pricing and nonconvex costs, unexpected regulatory changes that expand renewable deployment increase the uncertainty around the profits of investments in thermal generation plants. This uncertainty might preclude the realization of new investments necessary to guarantee the security of supply. The potential impact of the uncertainty around regulatory shocks on generation capacity was also identified at the inception of the spot pricing theory. As with any other market in equilibrium, in electricity generation "investments are made until the last MW of investment capacity 'earns' an expected stream of new income whose present value equals the incremental cost of investment" (Schweppe et al., 1987). However, in real systems, "unexpected future developments that upset the expected revenue streams [...] may result in imperfect revenue recovery". At the time, it was assumed that under spot pricing the magnitudes of such disruptions would be small assuming that capacity additions would take place in systems exhibiting a gradual growth of their net load. Yet, the magnitudes of the disruptions caused by the uncertainty in the total renewable capacity have been found to be non-negligible, violating that assumption. It is thus necessary to better understand the impact of these changes on the profitability of different technologies, and to introduce better ways for investors to value the actual risk to which plants are exposed so that better investment decisions are made.

If plans to deploy incentives for renewable generation are announced well in advance, and deployment occurs progressively with sufficient time for investors to respond, then investment decisions should take into account the differing operational regime that they will be exposed to in the presence of intermittent renewables. In this situation, it is appropriate that investors bear with the costs associated with failing to adapt to the announced future renewable capacity. On the other hand, if the regulatory decision leaves insufficient time for investors

to anticipate the change, underutilized assets can become partially or totally stranded, and a recovery procedure might be introduced by the regulator. The problem of shocks in the regulation pertaining renewable energy deployment is analogous to the shock emerging during the transition from a centralized planned system to a liberalized market [4].

1.3 Results from Optimal Marginal Pricing Theory

A comprehensive analysis of the economic principles underlying marginal pricing was first developed for electric power systems by Schweppe and his colleagues with their work on spot pricing of electricity (Schweppe et al., 1987). According to this theory, if generation is optimally dispatched, then marginal energy costs will exceed average operating costs, yielding revenues that will "exactly match operating and capital costs". Therefore, under ideal conditions, electricity marginal prices will induce the optimal level of investment, the optimal operation of the generation units and, in equilibrium, the profit received by generators would be equal to zero.

Pérez-Arriaga and Meseguer (Perez-Arriaga and Meseguer, 1997) expanded this theory by incorporating into the optimal prices the economic signals that generators and consumers must receive to meet short and long-term security of supply and other objectives as well as the impact of other type of constraints in order to maximize net social welfare. Optimal marginal prices must thus reflect: 1) the system marginal energy cost, 2) the system marginal reserves cost and 3) the system marginal cost for security of supply. Under this pricing scheme, if the total generation capacity of some technology is adapted, then the profit (i.e., the difference between revenues and operating and fixed costs) for this technology has to be exactly zero. They also showed that, under perfect competition, the solutions obtained by the following three distinct regulatory paradigms are equivalent:

1. The regulator as a single decision-maker.

2. A fully-competitive generation market.

[4]Details on the regulatory processes that took place in several power systems during the introduction of competition in the generation activity can be found in (Rothwell and Gomez, 2003)

3. A monopolistic utility under regulated supervision, whose cost function includes the cost of non-served energy.

Hence, the solution obtained through a decentralized individual welfare maximization by generators and consumers based on non-regulated prices is equivalent to the solution obtained through minimizing the total cost of a regulated monopoly. This result has been used to justify cost minimization formulations that include the cost of non-served energy (VOLL) in the analysis of perfectly competitive markets.

All these theoretical contributions are based on the interpretation of the first order conditions (FOCs) of the Lagrangian function formed with the system cost function (including investment and operation costs) and the system constraints (Perez-Arriaga, 1994). Therefore, the necessary and sufficient conditions to guarantee the optimality of the prices, and the investment and the operation solution obtained through this method are: 1) the differentiability and the convexity of the Lagrangian function, and 2) the convexity of the feasible set defined by the constraints (Floudas, 1995). In the context of power systems these conditions imply the following assumptions:

1. Investment decision variables are continuous (plants are not lumpy).

2. The cost structure of the plants is convex (startup costs are convex), as well as the utility function of the consumers.

3. No economies of scale (investment cost per MW of capacity is invariant of the total installed capacity).

4. There is a perfectly competitive market: there is no market power and there is perfect information.

5. Generators' remuneration comes entirely from the sale of electricity in the short-term market , as well as from providing reserves and adequacy services.

We revisit next the mathematical derivation of these results, in order to better understand the composition of optimal marginal prices.

1.3.1 The Regulator as a Single Decision-Maker: a Regulator's Model

The optimal investment and operation decisions in a system that is centrally-managed by a regulator maximize the net social benefit (NSB), which is equal to the total utility ($U(\cdot)$) received by consumers from using electricity, minus the cost of supplying that electricity. Therefore, the optimal operation and investment decisions are given by the solution of the following optimization problem (1.1)-(1.22).

$$\max_{x_{ij},\bar{x_i},D_j} NSB = \max_{x_{ij},\bar{x_i},D_j} \sum_{j\in\mathcal{J}}[U\left(D_j\right) - \sum_{i\in\mathcal{I}} VC(x_{ij})] - \sum_{i\in\mathcal{I}} \quad IC(\bar{x_i}) \tag{1.1}$$

$$\text{s.t.: } D_j - \sum_{i\in\mathcal{I}} x_{ij} \leqslant 0 \qquad \perp \mu_j \qquad \forall j \in \mathcal{J} \tag{1.2}$$

$$x_{ij} - \bar{x_i} \leqslant 0 \qquad \perp \psi_{ij} \qquad \forall i \in \mathcal{I}, j \in \mathcal{J} \tag{1.3}$$

$$\alpha_i\ \bar{x_i} - x_{ij} \leqslant 0 \qquad \perp \omega_{ij} \qquad \forall i \in \mathcal{I}, j \in \mathcal{J} \tag{1.4}$$

$$R(x_{ij}) = 0 \qquad \perp \zeta_{ij} \qquad \forall i \in \mathcal{I}, j \in \mathcal{J} \tag{1.5}$$

where $i \in \mathcal{I}$ and $j \in \mathcal{J}$ are the indices for power plants and time respectively; $x_{ij} \in \mathbb{R}_+$ is the decision variable for the energy output of power plant i at time j; $\bar{x_i} \in \mathbb{R}_+$ is the decision variable for the installed capacity at power plant i; $VC(\cdot)$ is the variable cost function; $IC(\cdot)$ is the investment cost function; D_j is the total demand at time j; $\alpha_i \in (0, 1)$ is the minimum output of plant i, expressed as a fraction of the plant's total capacity; and $R(\cdot)$ is a function representing other technical constraints of the plant that are common to the regulators and the generators problem.

The Lagrangian function of the regulator's problem becomes:

$$\mathcal{L} \equiv NSB^* = \sum_{j\in\mathcal{J}}[U(D_j) - \sum_{i\in\mathcal{I}} VC(x_{ij})] - \sum_{i\in\mathcal{I}} IC(\bar{x}_i) - \mu_j(D_j - \sum_{i\in\mathcal{I}} x_{ij})$$
$$-\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} \psi_{ij}(x_{ij} - \bar{x}_i) - \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} \omega_{ij}(\alpha\ \bar{x}_i - x_{ij}) - \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} \zeta_{ij}\ R(x_{ij}) \tag{1.6}$$

Taking the derivative with respect to each of the decision variables in the problem, we can derive the FOCs, (1.7) and (1.9).

$$\frac{\partial NSB^*}{\partial x_{ij}} = -\frac{\partial VC(x_{ij})}{\partial x_{ij}} + \mu_j - \psi_{ij} + \alpha\omega_{ij} - \frac{\partial R(x_{ij})}{\partial x_{ij}}\zeta_i = 0 \tag{1.7}$$

$$\frac{\partial NSB^*}{\partial \bar{x}_i} = -\frac{\partial IC(\bar{x}_i)}{\partial \bar{x}_i} + \sum_{j\in\mathcal{J}} \psi_{ij} - \sum_{j\in\mathcal{J}} \alpha\omega_{ij} = 0 \tag{1.8}$$

$$\frac{\partial NSB^*}{\partial D_j} = \frac{\partial U(D_j)}{\partial D_j} - \mu_j = 0 \tag{1.9}$$

If we solve for μ_j, we have that the optimality condition for the regulator's model is:

$$\frac{\partial U(D_j)}{\partial D_j} = \mu_j = \frac{\partial VC(x_{ij})}{\partial x_{ij}} + \psi_{ij} - \alpha\omega_{ij} + \frac{\partial R(x_{ij})}{\partial x_{ij}}\zeta_i \tag{1.10}$$

This equation shows that, in equilibrium, the marginal cost of supplying one more unit of electricity –which is a function of the variable cost and the dual values associated with the active constraints of the marginal plant–, has to be equal to the marginal utility that the demand perceives from consuming one more unit.

1.3.2 The Competitive Market Models

We now describe the models that describe the behavior of the system with perfect competition. Under this paradigm, we have two types of agents –consumers and generators– that decide upon the decision variables that they control with the objective of maximizing their own net benefit. Both consumers benefit, NCB, and generators benefit, NGB, are a function of a set of unregulated prices $\rho_j, j \in \mathcal{J}$ that are determined by the equilibrium

between supply and demand.

Consumers' Model

Consumers try to maximize their net benefit from electricity consumption.

$$\max_{D_j} NCB = \max_{D_j} \sum_{j \in \mathcal{J}} [U\,(D_j) - \rho_j\ D_j] = NCB^* \tag{1.11}$$

If we take the derivative with respect to the decision variable D_j, we have that the FOC of this model is:

$$\frac{dNCB^*}{dD_i} = \frac{dU(D_j)}{dD_j} - \rho_j = 0 \tag{1.12}$$

Generator's Model

Analogously to the consumer's model, generators try to maximize their net benefit from producing electricity, NGB. Note that in this model and in the consumer's model there is no demand balance equation enforcing that generation is equal to the total electricity demanded.

$$\max_{x_{ij}, \bar{x_i}} NGB = \max_{x_{ij}, \bar{x_i}} \sum_{j \in \mathcal{J}} [\rho_j \sum_{i \in \mathcal{I}} x_{ij} - \sum_{i \in \mathcal{I}} VC(x_{ij})] - \sum_{i \in \mathcal{I}}\ IC(\bar{x_i}) \tag{1.13}$$

$$\begin{aligned}
&\text{s.t.: } x_{ij} - \bar{x_i} \leqslant 0 && \perp \psi_{ij} && \forall i \in \mathcal{I}, j \in \mathcal{J} \\
&\qquad \alpha\ \bar{x_i} - x_{ij} \leqslant 0 && \perp \omega_{ij} && \forall i \in \mathcal{I}, j \in \mathcal{J} \\
&\qquad R(x_{ij}) = 0 && \perp \zeta_{ij} && \forall i \in \mathcal{I}, j \in \mathcal{J}
\end{aligned}$$

The Lagrangian function of this problem becomes:

$$\begin{aligned}
\mathcal{L} \equiv NGB^* = \sum_{j \in \mathcal{J}} [\rho_j \sum_{i \in \mathcal{I}} x_{ij} - \sum_{i \in \mathcal{I}} VC(x_{ij})] - \sum_{i \in \mathcal{I}} IC(\bar{x_i}) \\
- \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} \psi_{ij}(x_{ij} - \bar{x_i}) - \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} \omega_{ij}(\alpha\ \bar{x_i} - x_{ij}) - \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} \zeta_{ij}\ R(x_{ij})
\end{aligned} \tag{1.14}$$

If we now take the derivative of the Lagrangian function with respect to each of the decision variables of the generators, we have that the FOCs of the model are:

$$\frac{\partial NGB^*}{\partial x_{ij}} = \rho_j - \frac{\partial VC(x_{ij})}{\partial x_{ij}} - \psi_{ij} + \alpha\omega_{ij} - \frac{\partial R(x_{ij})}{\partial x_{ij}}\zeta_i = 0 \tag{1.15}$$

$$\frac{\partial NGB^*}{\partial \bar{x}_i} = -\frac{\partial IC(\bar{x}_i)}{\partial \bar{x}_i} + \sum_{j\in\mathcal{J}} \psi_{ij} - \sum_{j\in\mathcal{J}} \alpha\omega_{ij} = 0 \tag{1.16}$$

We can derive the optimality condition of the competitive market, taking the FOCs of the consumer's and the generator's model. This optimality condition implies that, in equilibrium, the marginal utility of demand is equal to the marginal cost of energy, as it was expected:

$$\frac{\partial U(D_j)}{\partial D_j} = \rho_j = \frac{\partial VC(x_{ij})}{\partial x_{ij}} + \psi_{ij} - \alpha\omega_{ij} + \frac{\partial R(x_{ij})}{\partial x_{ij}}\zeta_i = \mu_j \tag{1.17}$$

1.3.3 The Regulated Monopolistic Utility

In this model a monopolistic generation utility minimizes the supply cost, SC, resulting from total operation and investment costs.

$$\min_{x_{ij},\bar{x}_i} SC = \min_{x_{ij},\bar{x}_i} \sum_{j\in\mathcal{J}}\sum_{i\in\mathcal{I}} VC(x_{ij}) + \sum_{i\in\mathcal{I}} \quad IC(\bar{x}_i) \tag{1.18}$$

$$\text{s.t.: } D_j - \sum_{i\in\mathcal{I}} x_{ij} \leqslant 0 \qquad \perp \mu_j \qquad \forall j \in \mathcal{J} \tag{1.19}$$

$$x_{ij} - \bar{x}_i \leqslant 0 \qquad \perp \psi_{ij} \qquad \forall i \in \mathcal{I}, j \in \mathcal{J} \tag{1.20}$$

$$\alpha_i\ \bar{x}_i - x_{ij} \leqslant 0 \qquad \perp \omega_{ij} \qquad \forall i \in \mathcal{I}, j \in \mathcal{J} \tag{1.21}$$

$$R(x_{ij}) = 0 \qquad \perp \zeta_{ij} \qquad \forall i \in \mathcal{I}, j \in \mathcal{J} \tag{1.22}$$

Here non-served energy is treated as an additional technology with an investment cost equal to zero and a variable cost equal to the value of lost load (VOLL), which is the value that consumers put on electricity that is not consumed above the scarcity level.

The Lagrangian function of the regulated utility problem becomes:

$$\mathcal{L} \equiv SC^* = \sum_{j\in\mathcal{J}}\sum_{i\in\mathcal{I}} VC(x_{ij}) + \sum_{i\in\mathcal{I}} IC(\bar{x}_i) + \mu_j(D_j - \sum_{i\in\mathcal{I}} x_{ij})$$
$$+\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} \psi_{ij}(x_{ij} - \bar{x}_i) + \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} \omega_{ij}(\alpha\ \bar{x}_i - x_{ij}) + \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} \zeta_{ij}\ R(x_{ij}) \tag{1.23}$$

Taking the derivative with respect to each of the decision variables in the problem, we can derive the FOCs, (1.7) and (1.9).

$$\frac{\partial SC^*}{\partial x_{ij}} = \frac{\partial VC(x_{ij})}{\partial x_{ij}} - \mu_j + \psi_{ij} - \alpha\omega_{ij} + \frac{\partial R(x_{ij})}{\partial x_{ij}}\zeta_i = 0 \tag{1.24}$$
$$\frac{\partial SC^*}{\partial \bar{x}_i} = \frac{\partial IC(\bar{x}_i)}{\partial \bar{x}_i} - \sum_{j\in\mathcal{J}} \psi_{ij} + \sum_{j\in\mathcal{J}} \alpha\omega_{ij} = 0 \tag{1.25}$$

If we clear for μ_j with the FOCs, we have that the optimality condition for the regulated utility is:

$$\mu_j = \frac{\partial VC(x_{ij})}{\partial x_{ij}} + \psi_{ij} - \alpha\omega_{ij} + \frac{\partial R(x_{ij})}{\partial x_{ij}}\zeta_i \tag{1.26}$$

1.3.4 Key Results

It is straightforward to realize that the optimality conditions derived from the regulator's model (1.10) are equivalent to those derived from the models in the competitive market paradigm (1.17). Therefore, the optimal prices derived from the competitive market model, ρ_j, are equal to the Lagrangian multipliers of the demand balance constraint, μ_j, in both the regulator's and the competitive market's models. This result, together with other important results derived from optimal marginal electricity pricing are summarized as follows:

1. Under perfect competition, the solution of the regulator's model and the solution of the competitive market are equivalent.

2. In equilibrium, the profits of all units are exactly zero. Otherwise, new capacity would enter the market until equilibrium is reached.

3. The profits obtained by a technology whose capacity is optimized for a given mix of some other technologies, is exactly zero.

4. The existence of common or internal active constraints on installed capacity creates a positive or negative mismatch between revenues and cost of generators.

1.4 Impacts of Intermittent Renewables on Power Systems

This section provides a qualitative description of some of the operational and market impacts of intermittent renewables on electric power systems, some of which will be further quantified in the analyses in this thesis.

1.4.1 Renewables, Adaptation and Windfall Losses

The need to mitigate climate change through reducing anthropogenic green-house gas emissions and the absence of an international framework that regulates such emissions has urged many countries to mandate levels of renewable electricity deployment. This requirement is commonly achieved through price incentives (feed-in-tariffs) or through regulatory obligations (such as Renewable Portfolio Standards in the U.S.) (IEA, 2013), which might constitute a form of market distortion if the deployment of incentives does not leave time for investors in generation to anticipate to the new situation. Therefore, and given that renewable technologies can be installed much faster than other generation technologies, it is likely that rapid deployment of renewable generation will have implications on the dynamics of the market and that adaptation problems with the existing generation fleet will ensue. The resulting situation can be illustrated by the net load duration curve (NLDC)[5], showing how the net energy demanded is reduced as increasing quantities of wind and solar PV capacity

[5]The net load duration curve defines the probability distribution of the hourly difference between demand and renewable generation (i.e., the probability distribution of hourly energy that has to be supplied with conventional generators).

are introduced in the system (Figure 1-1). If we assume that initially, with no wind power or solar PV installed, the generation capacity in the system is perfectly adapted, the expected effects of a quick deployment of renewable capacity on the existing thermal capacity can be summarized as follows:

- In the short-term, there will be excess of capacity as new renewable units are pushed into an already adequately dimensioned system. This surplus of capacity would impose some negative consequences on units with high variable costs and high energy bids as these units are displaced by others bidding lower prices ('quantity effect'). This 'quantity effect' would in turn depress prices, further reducing the cash flows of all units in the system ('price effect').

- In the long-term, plants with higher capital costs could see their full-load hours reduced by larger shares of renewables as renewables increasingly reduce the baseload portion of the NLDC. Additionally, plants with high startup costs and plants with a high minimum output requirement would see their operating costs increase as a result of their exposure to the intermittency of renewable generation sources. In equilibrium, some of the plants with a high capital cost as well as plants that are 'inflexible would be decommissioned once their life expires, and substituted by 'flexible technologies that can follow the net load profile at a lower cost, as the generation fleet adapts to the new situation.

Figure 1-1 shows an example of the evolution of the NLDC with larger wind capacity (left panel) and solar PV capacity (right panel). This example was built upscaling wind and solar capacity factors to the given wind and solar capacities in each case (Holttinen et al., 2010). According to this illustration, the total energy that has to be supplied by conventional technologies will diminish with larger renewable shares. Yet, this figure does not include the variability of renewable resources, which also affects investment, commitment and dispatch decisions. Overall, the impact of net load changes on conventional generation units will be different depending on the units capex and opex values, and their flexibility characteristics.

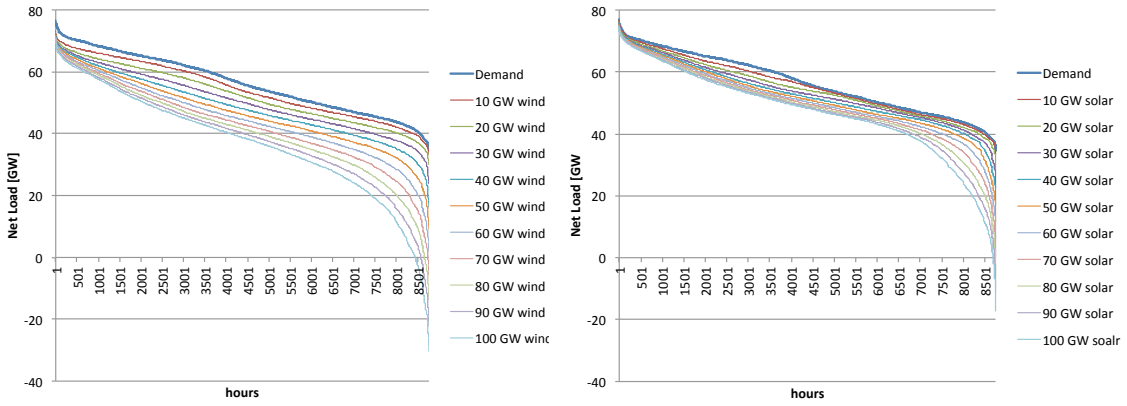


Figure 1-1: Net load duration curves for different wind power and solar PV penetration scenarios. Demand, wind and solar data correspond to the German System in 2008.

If changes in renewable capacity follow a deployment schedule that allows the other players in the market to anticipate future changes and adapt, the capacity mix in equilibrium will roughly consist of less plants with high capex, and more flexible plants that can adapt to the intermittency of renewables.

Conversely, without adaptation, it is expected that plants will see their capacity factor reduced as some of the energy is supplied by renewables, and see lower prices as a result of the ensuing excess of capacity. This effect can be characterized as windfall losses resulting from an unanticipated regulatory shock, to which the generation mix did not have time to adapt.

1.4.2 Operational Characteristics of Power Systems with Wind and Solar PV

Intermittency: Variability and Uncertainty

Variability and uncertainty have always been present in power systems: demand varies throughout the day and throughout seasons following consumers changing consumption of electricity, and generators must adjust their output accordingly to track this pattern. Additionally, demand is also uncertain by nature, even though the main drivers for demand are

very well known and it can be predicted very accurately. Similarly, power plant outages, hydro inflows, the possibility of a line tripping or other external factors contribute to increasing the uncertainty in the operation of power systems.

Although to some extent all power generation technologies can exhibit some variability and uncertainty for the reasons just described, wind and solar resources are extraordinarily variable and uncertain as they are fully dependent on instantaneous wind speed and insolation respectively, accouting for the sudden loss of very large plants. Many studies have shown how the spatial aggregation of wind turbines contributes to decreasing the system-level variability of wind output. Nevertheless, the variability of wind power even at a regional or national level is still high (Holttinen et al., 2008). In contrast to wind, solar power exhibits a more regular variability pattern, as solar availability is fundamentally driven by diurnal cycles, seasonal patterns and the presence of clouds and precipitation. Besides variability, wind and solar power availability are difficult to predict, although better forecasts can be achieved with proximity to real time. Therefore, wind and solar generation can both be considered as intermittent resources in the sense that intermittency accounts simultaneously for the non-controllable variability and the partial unpredictability of these resources (Perez-Arriaga, 2011).

The experience gained by independent system operators (ISOs) during the last decades has enabled them to cope with these uncertainties by introducing improvements in the management of the system, introducing redundancies through reliability criteria and deploying ancillary services. Ancillary services are "all non-energy services used to compensate for the variability and the uncertainty in the system in order to serve load reliably and keep system frequency stable" (Milligan et al., 2010). Among them, reserves can be defined as "real power capability that can be given or taken in the operating timeframe to assist in generation and load balance and frequency control" (Milligan et al., 2010). Table 1-2 presents a classification of types of reserves in terms of their response time and whether they are spinning (margin of capacity up or down from plants that are already producing) or non-spinning (margin of capacity up from plants that are not producing, or extra demand that can be called upon

request of the system operator). The names of the reserves used in this classification is based on the denomination used in the different north-american interconnections. For clarity, the European denomination is also provided in parenthesis.

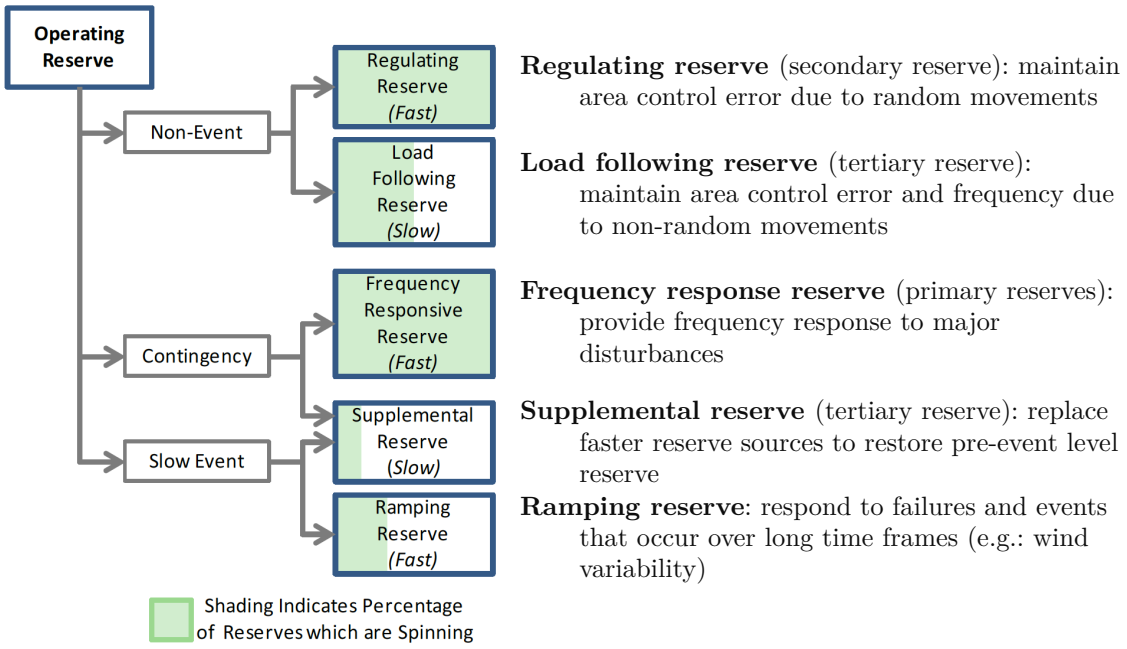


Figure 1-2: Illustration of reserves types with the European denomination in parenthesis. Adapted from (Milligan et al., 2010)

The response of power systems in the face of an increasing uncertainty from larger shares of wind and solar has been to redefine some of these reserve requirements (Holttinen et al., 2008)(Perez-Arriaga, 2011). Yet, larger renewable shares do not translate automatically into a proportional increase of reserve requirements. Many studies have shown the existence of a 'smoothing effect' from the spatial aggregation of intermittent resources, and many system operators currently use an 'aggregated' observation of the instantaneous renewable generation in the operation of the system for better efficiency. In addition to the aggregation effect, there are regions that might exhibit some correlation between demand, and wind and solar resources. Therefore, it is the net load (the difference between demand and renewable output) that ultimately matters in the operation and planning of power systems (Holttinen et al., 2010). The analyses in this thesis also takes this view and use net load as a joint

characterization of demand, wind, and solar resources.

Lastly, scarcity events extended for long periods of time occur with very low probability in systems with intermittent renewables. However, their impact on the system is far from negligible: they can provoke involuntary load curtailment or lead to the over-dimensioning of the generation capacity, both very costly outcomes. In order to cope with these events, systems could enable demand response mechanisms that reduce demand levels during times of resource scarcity or provide price incentives to activate back-up generators. This thesis does not explore the occurrence of these events, which should be analyzed on a region-by-region basis and by means of extensive historical data.

Cycling

The intermittency of renewable generation induces a more intense cycling regime on thermal power plants, as they have to start-up and shut-down more often in order to keep an instantaneous balance between demand and generation, or in order to meet reserve requirements (Perez-Arriaga and Batlle, 2012)(Holttinen et al., 2008). In particular, units that are high in the merit order like CCGTs or OCGTs are very exposed to renewable intermittency (given their high variable cost, they are the first being displaced by renewables) and they will quickly experience an increase in cycling. In some power systems, units considered as baseload (mostly coal units) are also providers of reserves and might see their part-load operation and cycling increased as a consequence of larger renewable shares (Troy et al., 2010).

The start-up and shut-down of a unit has a direct impact on the fuel cost and on the depreciation of the turbine. First, the startup of a thermal unit requires an extra amount of fuel to increase the temperature of the boiler to bring the unit to its minimum operational point. This amount of fuel required to startup the unit is a function of the time that the unit has been offline. Second, cycling operations moves the unit to operate more frequently within less efficient regimes in the heat rate curve, requiring more fuel to generate electricity than in the optimal generation region. Third, the number of starts is a major driver for the

wear and tear of the turbine, and a more intense cycling regime will accelerate the need to major turbine overhauls, dramatically affecting the profitability of the unit. In particular, the long-term service agreements between the plant and the original equipment manufacturer, establish the needs for inspections and maintenance services, as well as for major overhauls, which are in the range of $20-60M each (Rodilla et al., 2013). Therefore, given its economic importance, it is essential to include in detail the costs of starting up in models used to analyze investment decisions and markets, such as capacity expansion models. Figure 1-3 represents three different maintenance interval functions of a gas turbine, corresponding to three possible ways of accounting for its firing hours and the number of startups (Option A, maintenance is determined by the first threshold criteria reached; Option B, each start cycle has an equivalent number of operating fire hours; and Option C, maintenance is determined by a function of both firing hours and cycles performed).

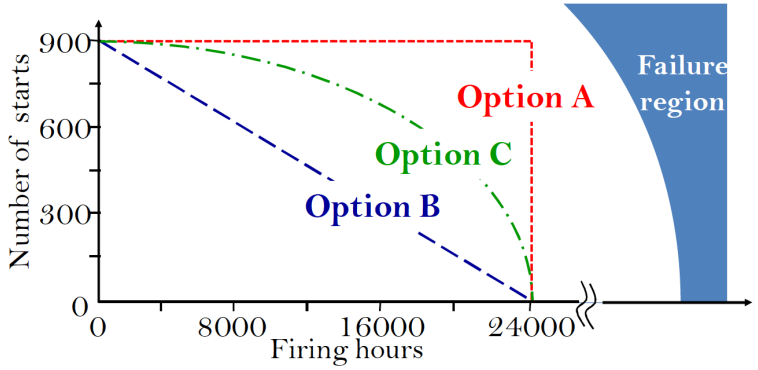


Figure 1-3: Baseline functions for maintenance interval (Rodilla et al., 2013)

In the analyses performed in this thesis, startup costs are modeled as a single quantity that includes both fuel costs and major overhauls, without considering explicitly the conditions included in maintenance contracts, like those in Figure 1-3.

Flexibility

In general, it has been acknowledged that systems aiming at introducing large amounts of intermittent renewables will be required to be more flexible (Lannoye et al., 2011) (Lannoye et al., 2010) (IEA, 2011) (Holttinen et al., 2008) (Perez-Arriaga, 2011). At a system level,

flexibility can be defined as the capacity to adapt generation and demand in order to integrate intermittent resources. This definition is consistent with other definitions found in the literature (IEA, 2011).

There are multiple elements that can be introduced to enhance the flexility of a power system, mainly: storage, demand side response (DSR), interconnections with neighboring areas and 'flexible' power plants. Each of these 'flexibility options' have different implications on the operation of power systems, and present different advantages relative to one another. The International Energy Agency (IEA) has extensively explored the impact of each of these options on the operation of the system, as well as their economic tradeoffs, using the same modeling approach as the analyses in this thesis (IMRES) (IEA, 2014).

Focusing on thermal generation, power systems with renewables will require a larger number of flexible generating units that balance renewable electricity variability, some of which will operate for fewer hours of the year (DENA, 2010)(Denholm and Margolis, 2007)(Denholm et al., 2010)(IEA, 2011)(Lannoye et al., 2011)(Milligan et al., 2010)(MIT, 2011)(Perez-Arriaga, 2011). Operational flexibility as an attribute associated with power generators entails some specific characteristics: low start-up costs (ability to cycle), highly adjustable output rate (low minimum operating point) and high ramping rate (fast ramp-up capacity) (Lannoye et al., 2010) (Perez-Arriaga, 2011). Some research efforts have been lately geared towards identifying which of all of the attributes associated with flexible thermal plants are mostly needed in systems with intermittent renewables, in order to create new services that can be priced to compensate flexible plants for their extra fixed costs compared to conventional generation technologies. Most of these studies have shown that the primary scarce attribute is fast ramping capability (ramping that occurs in a timeframe of minutes) (Lannoye et al., 2010) and some power systems have started to introduce mechanisms that remunerate plants providing this service (California ISO, 2011).

As this thesis focuses on day-ahead wholesale markets, the analysis of the intra-hourly variability of the net load (what happens in timeframes of minutes) is beyond its scope and, therefore, the flexibility attributes included in the analyses of this thesis are those that are

salient in timeframes greater or equal than one hour.

1.4.3 Wholesale Electricity Markets and Intermittent Renewables

There are many open questions regarding how intermittent renewables affect electricity wholesale markets. During transitory stages, since renewables have zero variable cost, some systems with large penetration of intermittent renewables have experienced a decrease in prices as the units with high variable cost are partially or totally displaced by renewable generation. This price effect might be partially counteracted if wholesale prices account for startup costs as will be shown in Chapter 3.

An additional contributor to decreasing prices is the curtailment of renewables, resulting in zero price hours whenever curtailment occurs. At moderate penetration levels, renewable curtailment occurs for two primary reasons: 1) a lack of available transmission to integrate part of the renewable energy generated; and 2) situations where renewable generation exceeds load and the excess cannot be exported to other balancing areas due to transmission constraints (Fink et al., 2009). At high penetration levels, renewable curtailment adds a degree of freedom to the operation of the system, and it can be used to avoid violating ramping constraints, reserves constraints, or to avoid the cost of shutting down a unit and having to start it up again. Some markets introduce a priority dispatch condition as an incentive for renewable generation, by which renewables cannot be curtailed unless they are jeopardizing the safe operation of the system. Nevertheless, this condition introduces an inefficiency in the system, as the possibility of curtailment can reduce the operational cost of the system. Hence, all the analyses in this thesis include renewable curtailment as an option that can be used to reduce the total system cost.

In equilibrium, it is not clear what the evolution of electricity prices would be, as there are other factors such as the formats of bidding, or changes in fuel prices that could shift the capacity mix and the prices in one direction or the other.

In addition to the characteristics mentioned, an effect frequently observed in markets with intermittent generation is an enhanced volatility, with an increasing occurrence of high

price spikes and numerous occurrences of zero or even negative prices[6] (Perez-Arriaga and Batlle, 2012). The impacts resulting from all these market effects on the value of generating assets are not fully clear yet. This thesis addresses these issues by focusing on the profits obtained by generators under different renewable penetration scenarios and different formats of bidding, each exhibiting different magnitudes of volatility and renewable curtailment.

1.4.4 Taxonomy of the Impacts of Intermittent Renewables on Electricity Markets

Table 1.1 summarizes the main factors that have been found to have an impact on the profitability of thermal generating units and, ultimately, on the risk perceived by investors in electricity generation. The factors covered in this thesis have been marked in bold.

This thesis the impacts of large-scale deployment levels of intermittent renewables on electricity markets, focusing on the impact of pricing rules and unanticipated regulatory decisions that increase renewable capacity, on the wholesale electricity price, and on the profits made by generating technologies. In addition, as individual plants' profit calculations are based on hourly prices, they already account for the effect of price volatility on plants' profits.

Since the trade of electricity in different time scales is typically referenced to the prices of the day-ahead market[7], this thesis focuses exclusively on this market, leaving secondary and reserve markets beyond the scope of this research. Likewise, the study of the impact of rare scarcity events, external factors, and intra-hourly ramps on electricity markets requires exhaustive statistical analysis and sensitivity analysis with data specific to the region of study, and are also beyond the scope of this research.

[6] In power systems where renewables do not enjoy priority dispatch while receiving feed-in tariffs, renewables might see an incentive to bid negative prices in the market to ensure being dispatched. Yet, feed-in-tariffs and other incentives are not considered in this thesis and it will be assumed that renewable plants always bid zero prices.

[7] The day-ahead electricity market is a forward market where participants submit quantity-price bids, sometimes together with other conditions, for the 24 hours of the following day. Typically, bids in this market are submitted between 30 and 36 hrs before the delivery time

Table 1.1: Summary of the main factors affecting the profitability of thermal units.

Factor	Origin	Analysis
1. Price volatility	Zero or negative prices and price spikes	Obtain profit distribution with and without renewables
2. Scarcity events	Rare periods of renewable resource scarcity	Simulation of weeks including rare events found in long-term data series analysis
3. Regulatory decisions	Rapid deployment and/or regulatory uncertainty on renewable deployment	Obtain profit distribution w/ renewables in a system designed for less or no renewables
4. External factors	CO2 price, different gas prices	Sensitivity analysis under different conditions
5. State of the market	Correlation between capital markets and conditions affecting generation assets	Calculation of stochastic discount factors and risk neutral discounting
6. Pricing rules	Ability of pricing rules to recover costs from an increased cycling regime	Obtain the profit distribution with different start-up costs and different pricing rules
7. Ramps	Increase in the variability of the net load	Analysis of net load ramps at different time-steps

1.5 Methodology

1.5.1 Electricity Market Models

As shown above, under perfect competition, and with a convex feasible set and cost functions, economic theory demonstrates that centralized planning models that minimize investment and operation costs and include the cost of non-served energy (VOLL) provide the same solution as market models where the objective is to maximize generators' and consumers' total profits (Green, 2000)(Perez-Arriaga, 1994)(Perez-Arriaga and Meseguer, 1997). This equivalence allows for the use of traditional cost-minimization models to simulate specific results pertaining to situations of perfect competition.

In highly concentrated markets, firms try to exert market power and the size of firms and the prices diverge from those in perfectly competitive markets (Stoft, 2002). Under these conditions, models that assume perfect competition do not accurately represent the behavior

of the market in question, and therefore the strategic behavior of the different players has to be considered (Hobbs, 1995). One of the most common approaches to represent perfectly and imperfectly competitive markets is using equilibrium models. The solution of equilibrium models involves solving a system of equations built upon the KKT optimality conditions of the original problem (Ventosa et al., 2005). Among these conditions, the *complementary slackness* condition is that at an equilibrium point, the product of the value of the constraint at that point and its dual variable equal zero. Nevertheless, one necessary condition for this powerful result to hold is that the set of feasible solutions over which the optimization takes place is convex (Floudas, 1995), which is rarely the case for power systems with intermittent renewables and nonconvex costs. Therefore, this thesis assumes perfect competition in order to be able to account for nonconvex costs and focus exclusively on the effects of intermittent renewables on electricity markets.

In power systems, generation capacity expansion models have traditionally been used, among others, for investment planning, to analyze how plants can recover both fixed and variable costs, and to assess the environmental impact of electricity generation. Some of these models, such as those based on screening curves and the load duration curve, are useful to highlight the economic tradeoffs between the fixed costs and the variable costs of the available technologies in the system (Turvey and Anderson, 1977). Other more sophisticated models, also based on the load duration curve, incorporate more advanced techniques to account for the operational uncertainty of the units into the dispatch process in order to assess the reliability of the system (EPRI, 1980).The operational characteristics of power systems with a large amount of intermittent renewables bring additional costs to the system from starting up thermal units and from the activation of other constraints, as power plants need to ramp up fast and cycle more often to balance the intermittency of wind and solar PV (Lannoye et al., 2010). To estimate the additional costs requires investment models that incorporate some of the operational constraints that are not included in traditional capacity expansion models. Power system models that account for start up costs and operational constraints are called unit commitment models. These models choose the commitment state

of individual plants and dispatch decisions, accounting for the operational constraints of the plants in the system to find a technically feasible solution. Capacity expansion models that include the operational details that represent the effects of renewable electricity are computationally challenging. First, the use of binary variables to represent operation and investment decisions create a domain that is a combinatorial space, demanding sophisticated solution methods (Bertsimas and Tsitsiklis, 1997)(Bertsimas and Weismantel, 2005). Second, the combined variability of renewables and demand is such that the net load cannot be realistically represented anymore with a small number of blocks (Turvey and Anderson, 1977)(Holttinen et al., 2008). This increases the dimensionality of the problem making it computationally intractable for a very large number of load periods. Dimensionality problems are usually tackled with conventional stochastic programming (SD) decomposition techniques (Geoffrion, 1972)(Birge and Louveaux, 1997). However, adaptations of these techniques to nonconvex problems such as this one are only valid when the nonconvexity lies in the function being optimized and not in the feasible set defined by the constraints and the decision space (Li et al., 2011).

Recently, efforts have been made to account for more detailed operational constraints in perfect competition models in a computationally tractable way (Bushnell, 2010)(de Jonghe et al., 2011)(Ma et al., 2013)(Palmintier and Webster, 2013)(de Sisternes, 2011)(Batlle and Rodilla, 2013). The model developed in the context of this thesis has been inspired by some of these recent efforts to include the operational details necessary to account for the effects of renewable electricity in real-sized market investment models (Palmintier, 2012)(Palmintier and Webster, 2013). This model has been labeled '*Investment Model for Renewable Electricity Systems (IMRES)*' and consists of a cost minimization capacity expansion model –perfect competition, socially optimal– with unit commitment constraints embedded. IMRES has a two-component structure: in the first component, individual plant building decisions are made; in the second component, a unit commitment model decides on the short-term operation of the power plants that have been built, while meeting all operational constraints. The high-dimensionality problem is addressed by limiting the unit commitment component

of the model to a number of representative weeks that optimally represent the total energy demanded in the system (de Sisternes, 2013). In the next section it will be shown that this representation provides also a good approximation of the chrnological variability of the net load and the operation of thermal plants in the presence of intermittent renewables, although these parameters are not explicitly reflected in the approximation method.

Unlike other models that focus exclusively on cost, IMRES has been designed to model electricity markets –in terms of both costs and prices– in power systems with intermittent renewables. The design requirements of IMRES are threefold: first, hourly resolution to simulate wholesale prices; second, constraint that couple time periods in chronological order to reflect on investment decisions the extra costs derived from the operation of plants in a system with large shares of intermittent renewables; and third, individual treatment of plants to simulate profit results. IMRES' two-component structure of capacity expansion with unit commitment constraints reflects the impact of a more intense operating regime on investment decisions, which are represented by binary variables. Additionally, the net load approximation using whole weeks, simplifies the dimensionality of the problem while accounting for the chronological variability of the net load and the correlation between renewables and demand. The 'market simulator' component of IMRES, in which wholesale prices are derived, was calibrated to approximate to the expected theoretical result in equilibrium (i.e., total revenues equal total costs), while taking into account the divergences produced by the nonconvexities included in the model.

Throughout the analyses carried out in this thesis, IMRES is used to simulate the impacts of different penetration levels of renewables, in terms of the optimal capacity mix, and the capacity factor, the energy contribution, the cycling regime and the profits obtained by each thermal technology present in the system. In addition, IMRES is used to explore the impact of different bidding mechanisms (complex and simple bids) on the total cost recovery of units and on the cost paid by consumers.

Given the nonconvexities in the cost structure of the units and the changes in remuneration introduced by the different bidding mechanisms, the hourly price solution provided

by IMRES might not support a market equilibrium solution (i.e, some technologies might appear as under or over-compensated, and a different number of plants would have been installed in reality). Therefore, a metaheuristic approach is presented to approximate the solution to market equilibrium. Metaheuristic models combine optimization techniques with heuristic rules that improve candidate solutions. In particular, this model will calculate the expected revenues of each individual plant using a standard unit commitment model, and it proposes new candidate solutions according to a set of heuristic rules that reduce the number of plants from technologies that incur losses and increase the number of plants of technologies making large profits. These rules ensure that all the plants proposed in each solution obtain a positive profit and that no new entrant would make a profit in the new solution. Metaheuristic models such as the one proposed in this thesis allow us to search in a high-dimensional space, but their main disadvantage is that the equilibrium solutions produced by these models are not guaranteed to be global optima.

1.5.2 Valuation Models Accounting for Risk

Investors in electricity generation assets turn to various valuation techniques to discount future cash flows generated by the assets for time and risk, in order to obtain their present value to be used in strategic decisions. Most non-financial firms use the capital asset pricing model (CAPM) to derive a price for the risk that a prospective investment decision will entail. The principal risk metric used in this exercise is the beta of the asset, which measures how much on average the value of the asset will change given a 1% change in the market index[8](Brealey et al., 2008). The risk premium derived with this beta added to the risk-free rate of return is then used to discount the future cash flows of the asset, performing what is commonly known as risk-adjusted discounting. The main feature of this asset valuation method is that it assumes that the assets' risk is directly proportional to the market risk, ignoring other more complex relationships. In particular, in power systems there are caps on prices, operational constraints and regulatory changes –such as the unexpected deployment

[8]Here the market index represents the general risk of the market and it can be calculated as a weighted average of all the assets in the market

of renewable generation– that introduce a series of nonlinear and time-varying relationships between the state of the market and the cash flows obtained by a power plant.

This thesis proposes an algorithm based on stochastic discounting that makes explicit the relationship between the underlying market risk and the cash flows generated by the asset in order to accurately value risk. The algorithm proposed is divided into four steps: 1) the macro perspective step, modeling the evolution of the market index (our chosen risk factor), looking at how market risk (the macroeconomic risk) is structured, and deriving stochastic discount factors for each of the states considered; 2) the micro perspective step, modeling the relationship between the market index and the cash flows produced by the generation asset, and defining state-contingent cash flows; 3) the valuation step, where cash flows are discounted for time and risk to obtain the present value of the asset; and 4) a last step, calculating useful parameters such as an implied discount rate of the asset valued (i.e., a discount rate that would yield the same present value of the asset if expected cash flows are discounted at that rate).

1.6 Thesis Research Question and Contributions

The main hypothesis in this thesis is that the conditions introduced by a large penetration of intermittent renewables might prevent thermal plants from recovering costs, making investors in power plants to diverge in behavior from the welfare maximizer level of investment. Therefore, the main research question can be formulated as follows:

> What are the risk implications of a large penetration level of intermittent renewables on investments in electricity generation?

In particular, this question can be subdivided in the following five questions:

1. How do different penetration levels of intermittent renewables affect the optimal capacity mix; and the capacity factor, the energy contribution, the startup regime and the profits of individual units?

2. How do unexpected penetration levels of intermittent renewables impact the cost recovery of thermal generators in the wholesale electricity market and the cost paid by consumers?

3. How do different bidding rules impact the cost recovery of thermal generators in the wholesale electricity market and the cost paid by consumers?

4. How do parameters characteristic of the system such as the VOLL, the size of the system and the startup cost affect estimates of cost-recovery by plants?

5. How should investors value the risk of new investments in generation plants in order to make better investment decisions?

The goal of this thesis is to elucidate the impacts of intermittent renewables on electricity markets and, in particular, explain how investments in thermal technologies are affected by a large penetration of renewables. Specifically, this thesis will focus on: 1) explaining the effects in a system that is adapted for certain renewable capacity and also in a system that is designed for a zero renewables, but operated with a non-zero renewable capacity; 2) explain the effects of bidding rules on recovering operational and fixed costs of thermal power plants, and the impact on the cost paid by consumers; and 3) explain how risk should be accounted for in investment decisions.

Methodologically, this thesis presents a new capacity expansion model with unit commitment constraints, capable of analyzing the market effects of real-size power systems with large amounts of renewables. It also proposes a heuristic method with which to derive a market equilibrium solution in situations where the cost minimization solution does not support equilibrium. In addition to this, this thesis introduces a method based on stochastic discounting and Monte Carlo simulation to value the risk of new generation assets.

1.7 Thesis Structure

The thesis is structured as follows.

Chapter 2 describes the capacity expansion model with unit commitment constraints used in this research to explore the impacts of renewables on electricity markets. It also describes the heuristic method used to select the weeks used in the unit commitment component of the model, and shows how four weeks can approximate well the commitment results derived from the full one-year series.

Chapter 3 analyzes the impact of renewable generation on the cost recovery of thermal generators. In particular, it explores the effects of the unanticipated deployment of renewables and the impact of three different bidding rules (side payments, price uplift and simple bids) on plants' profits and the cost paid by consumers. In addition, this chapter summarizes the impact of renewables on the optimal capacity mix, and the operational characteristics of the generation technologies considered. It also analyzes the impact of some system parameters such as the size of the units, the VOLL and the startup cost on profit results.

Chapter 4 presents the issue of accounting for the value of risk in electricity generation investments and explains with an example the application of the Monte Carlo method to the stochastic discounting methodology for the valuation of electricity generation assets under uncertainty.

Finally, Chapter 5 summarizes the conclusions and the regulatory implications of the results found in the thesis.

Chapter 2

Methodology

This chapter is an adaptation of the following papers:

- *Investment Model for Renewable Electricity Systems (IMRES): an Electricity Generation Capacity Expansion Formulation with Unit Commitment Constraints*, MIT Center for Energy and Environmental Policy Research, CEEPR WP 2013-16, Cambridge, MA (de Sisternes, 2013)
- *Optimal Selection of Sample Weeks for Approximating the Net Load in Generation Planning Problems*, ESD Working Paper Series, ESD WP 2013-03, Cambridge, MA (de Sisternes and Webster, 2013)

2.1 Introduction

The ongoing large-scale introduction of intermittent renewable energy resources in electric power systems has substantially increased the variability and the uncertainty of the net load, the difference between the actual load and the variable energy output, also known as residual demand (Perez-Arriaga, 2011). Under this new paradigm, resources that can respond quickly and balance the variability of the net load throughout a prolonged period of time need to be deployed in the system (Milligan et al., 2010)(Morales et al., 2009)(Holttinen et al., 2008). These resources typically have low startup and variable costs, and provide

what is generally referred to as *operational flexibility* or, simply, *flexibility* (although there is no common definition for it) (Perez-Arriaga, 2011)(IEA, 2011)(Lannoye et al., 2011). Quantifying the operational flexibility needs in power systems with renewables requires using long-term investment models that account for the cost of deploying these resources, as well as including sufficient operational detail to account for the cycling costs and operational effects resulting from the additional variability introduced (Lannoye et al., 2011)(de Jonghe et al., 2011)(Palmintier and Webster, 2013)(IEA, 2014)(Ma et al., 2013). Additionally, the detailed analysis of electricity markets involves the derivation of wholesale prices, and a model structure with decision variables particular to individual plants, to enable the analysis of each plants' profits.

Therefore, the elements required for a capacity expansion model that assess the impacts of intermittent renewables on electricity markets –in terms of both costs and prices– are three:

1. Hourly resolution to allow the simulation of wholesale prices and the implementation of unit commitment constraints.

2. Unit commitment constraints that couple variables in chronological demand periods in order to reflect the additional costs derived from the distinct operation of plants in a system with large shares of intermittent renewables.

3. Individual treatment of plants (investment and operating decision variables particular to each plant) to simulate the profits obtained by individual units.

Capacity expansion problems are typically used to determine the optimal generation capacity mix and the level of reserves to supply electricity reliably to meet a given demand profile represented by a load duration curve (LDC). However, increased variability in the system causes a more intense cycling regime in thermal units, whose additional fuel and O&M costs associated with starting up is not captured in the classic version of the problem. Traditional capacity expansion models such as screening curves models (Turvey and Anderson, 1977) primarily capture the economic trade-off between generating technologies with

a high capital cost and a low variable costs, and technologies with lower capital cost but high variable cost, without accounting for other important factors such as start-up costs, the indivisibility of units, minimum stable output levels, ramp limits and reserve needs. Besides the simple versions of screening curves, there are other more sophisticated models based on load duration curves that account for reserve margins and reliability criteria (EPRI, 1980), but that do not include the commitment detail and the chronological ordering necessary to reflect the impact of renewables on the operation of the units.

Conversely, unit commitment (UC) problems represent the hour-by-hour dynamics of the system solving for commitment states, start-up and shutting-down of plants for a given capacity mix, but do not consider new investment decisions (Baldick, 1995). In addition to including the technical constraints in the system, UC problems for systems with high penetration of intermittent renewables, need to account for the additional uncertainty brought in by renewable resources. In previous studies, some authors have introduced methods that address some of the classic uncertainties present in power systems. For instance, the uncertain availability of plants has been considered through probabilistic reserve determination (Dilon et al., 1978), and uncertainty in demand (Takriti et al., 1996) and stochastic hydro-inflows in hydro-thermal systems (Pereira and Pinto, 1985) through stochastic programming. More recently, stochastic programming approaches (Morales et al., 2009)(Tuohy et al., 2009)(Papavasiliou et al., 2011) and stronger formulations of the problem (Ostrowski et al., 2012) have been proposed to more effectively and efficiently model the uncertainty posed by renewables, given an existing capacity mix.

Models that assess the need for flexibility for long-term planning in a system with renewables must simultaneously include capital investment decisions, operational decisions, and the uncertainty and variability from intermittent renewables. Recently, formulations based on including capacity decisions within existing simplified UC problems have gained momentum, as a means to assess operational flexibility needs (de Jonghe et al., 2011)(Palmintier and Webster, 2013)(Ma et al., 2013). These formulations minimize the total system cost (the sum of the fixed and variable costs) over one year, by taking into account the annualized

capital cost of the units and the total variable cost of a one-year unit commitment problem, including the main operational constraints. These models are computationally very difficult to solve as each potential unit in the system has on the order of thousands of binary commitment variables associated, in addition to the constraints on this variables. For instance, to make commitment and generation decisions from a pool of 300 generating units, one would need to solve a problem with over five million binary variables (counting each units commitment and start-up decisions over the 8,760 hours of a year). Yet, the high dimensionality of these models cannot be addressed with a straightforward application of stochastic optimization techniques, as binary commitment variables would be part of the subproblem in the classic Benders decomposition formulation, which impedes closing the duality gap (Birge and Louveaux, 1997).

Palmintier and Webster propose a method that incorporates operational flexibility into long-term capacity expansion planning, based on a capacity expansion formulation with unit commitment for a full 8 760 hour load profile (Palmintier and Webster, 2011)(Palmintier and Webster, 2013). This method is founded on an integrated clustering approach for the unit commitment, that combines together all units from the same technology into a single cluster. This approach improves the computational tractability of the model, but it is not suitable a priori for analyzing the profitability of individual units as they are bundled together in clusters.

Instead, Morales et al. point out that improving the computational tractability of models that assess the need for reserves in systems with renewables might require an appropriate bundling of hourly wind-related values and a reduced but representative set of renewable generation scenarios (Morales et al., 2009). Accordingly, Some authors have proposed simplified versions of such models, all of which select a number of representative weeks to construct an approximate load profile of the systems yearly demand (de Jonghe et al., 2011)(Ma et al., 2013) (Papavasiliou et al., 2011). De Jonghe et al account for the chronologic sequence of hourly load levels, including some operational constraints, but their approach decides upon technologies and not upon individual plants, which impedes accounting for startup costs and

the minimum up and down times of the plants (de Jonghe et al., 2011). Ma et al. based their selection of representative weeks on seasonality and well-known demand patterns, which does not reflect accurately the total amount of energy demanded in the system over the year (Ma et al., 2013). Conversely, Papavasiliou et al. proposed a method that optimally determines the weights of a pre-selected number of representative wind power scenarios (Papavasiliou et al., 2011). However, in a system with a large penetration of renewables, a sound method that selects a small number of representative weeks has to simultaneously provide a good account of the load duration curve while including potential correlations between the load and renewable output. Therefore, this thesis addresses the dimensionality challenge by optimally selecting a number of weeks from a one-year-horizon hourly load series, given the hourly availability of wind and solar power over the same timeframe, and taking into account the correlation between demand and renewable resources.

There are other approaches that are not strictly based on optimization, connecting long-term capacity expansion planning models with a detailed representation of the cycling costs of the plants. One example of these models is 'LEEMA', a model developed by Batlle and Rodilla that incorporates maintenance costs and major overhauls to the traditional screening curves model (Batlle and Rodilla, 2013). LEEMA is based on a heuristic optimization that calculates the production profile and the total cost that a 1MW fictitious plant of some certain technology will have if it supplies each of the 1MW horizontal slices in which the sequential hourly demand is divided (loading points). Capacity decisions are subsequently made attending to the cost functions resulting from assuming that the different production profiles are supplied by each thermal technology considered. Although hourly prices can be derived from this model, its heuristic nature would make it difficult to calibrate so that, in equilibrium, the profits obtained by all units approximate the expected theoretical result of zero profits.

This chapter describes the *Investment Model for Renewable Electricity Systems* (IMRES), a model based on a selection method that accounts for the correlation between demand and renewable availability, which aims at determining the minimum cost electricity generation

capacity mix and the profits obtained by individual units in systems with a high penetration of intermittent renewable resources. IMRES uses a capacity expansion formulation with embedded unit commitment constraints, integrating the operational dynamics induced by a high net load variability characteristic of this type of power systems. The main advantage of this formulation is that its objective function accounts for capital costs, variable costs, as well as the costs associated with a more intense cycling regime. In addition, the model also includes commitment constraints that couple together consecutive demand periods and relate the technical characteristics of thermal units to total system cost and capacity decisions. The dimensionality challenge is addressed by optimally selecting a number of weeks from a one-year-horizon hourly load series, given the hourly availability of wind and solar power over the same timeframe, and taking into account the correlation between demand and renewable resources.

From a centralized planning perspective, IMRES can help to determine the future investments needed to supply a future electricity demand at minimum cost. In the context of liberalized markets, IMRES can be used by regulators for *indicative energy planning* (Perez-Arriaga and Linares, 2008) in order to establish a long-term vision of where efficient markets should evolve, and to have an estimation of the impact of renewables and market rules on the profitability of generating units. The method presented with IMRES combines the economic assessment performed by classic approaches with the techno-economic analysis of unit commitment models, allowing a detailed study of the impact of technical constraints on the system cost and the profit made by individual units.

IMRES can be viewed as a two component model (Table 2.1): the primary component decides which power plants to build; and the secondary component accounts for the operational decisions at the power plants. In its original form, renewables and storage capacity are taken as parameters, while capacity and operational decisions concerning the thermal capacity mix are treated as variables. However, as will be shown below, IMRES also allows renewables and storage capacity to be treated as decision variables, at the expense of increased computational complexity and computation time.

Minimize	Investment costs + Operational costs	(2.45)

s.t.:	operate-if-built coupling constraint	(3.17)
	demand balance equation	(2.3)
	unit commitment constraints	(3.5-3.12)
	renewable energy and emissions constraints	(2.12-3.15)
	storage constraints	(2.16-2.20)
	demand-side management constraints	(2.21-2.25)
	reserves constraints	(2.26-2.32)
	non-negativity/binary constraints	(3.18-2.42)

Table 2.1: IMRES' General Structure and Reference to Equations

The time interval evaluated in IMRES is one year, divided into one-hour periods and representing a future year (e.g.: in 2050). In this sense, IMRES is a *static* model because its objective is not to determine when investments should take place over time, but rather to produce a snapshot of the minimum cost generation capacity mix under some pre-specified future conditions.

This chapter begins by presenting the mathematical formulation of the model, with a detailed description of the unit commitment constraints used. Second, it shows how the size of the problem posed by selecting weeks from a one-year data series increases with the number of weeks selected. Third, it proposes metrics to test the quality of the approximation, arguing that a small number of weeks is sufficient to represent the net load duration curve with negligible error. Fourth, it explores the trade-off between the added complexity of including the peak net load day in the problem and achieving a more accurate representation of the hours with non-served energy (NSE), which is critical for determining the economic feasibility of peaking units. Fifth, it demonstrates the improvement in accuracy from accounting for the correlation between demand and renewables, using a selection based on the net load duration curve, as opposed to a selection based on seasons or based on the system's load duration curve. Lastly, it validates the applicability of this method to power systems with a high penetration of renewables by comparing the error values obtained in two different

regions.

2.2 Notation

2.2.1 Indices and Sets

Table 2.2: IMRES' Indices and Sets

$i \in \mathcal{I}$, where $\mathcal{I}$ is the set of generating units that can be potentially built
$j \in \mathcal{J}$, where $\mathcal{J}$ is the set of hours in the data series
$j' \in \mathcal{J}$, where $\mathcal{J}$ is the set of hours in the data series
$k \in \mathcal{K}$, where $\mathcal{K}$ is the set of generation technologies
$\mathcal{W} \subset \mathcal{I}$, where $\mathcal{W}$ is the subset of wind units
$\mathcal{S} \subset \mathcal{I}$, where $\mathcal{S}$ is the subset of solar photovoltaic units
$\mathcal{T} \subset \mathcal{I}$, where $\mathcal{T}$ is the subset of thermal power units (nuclear, coal, CCGTs and OCGTs)
$\mathcal{T}^k \subset \mathcal{I}$, where $\mathcal{T}$ is the subset of thermal units of technology k
$\mathcal{G} \subset \mathcal{I}$, where $\mathcal{G}$ is the subset of gas peaking units (OCGTs)
$\mathcal{N} \subset \mathcal{I}$, where $\mathcal{N}$ is the subset of nuclear units

2.2.2 Variables

Table 2.3: IMRES' Variables

$y_i \in \{0, 1\}$	building decision for power plant i
$\boldsymbol{y}^*$	optimal solution for capacity expansion (individual building decisions)
$x_{ij} \in \mathbb{R}_+$	output power of plant i during hour j
$u_{ij} \in \{0, 1\}$	commitment state of power plant i during hour j
$z_{ij} \in \{0, 1\}$	start-up decision of power plant i at hour j
$v_{ij} \in \mathbb{R}_+$	shut-down decision of power plant i at hour j
$w_{ij} \in \mathbb{R}_+$	output power over minimum output of plant i during hour j
$f_j \in \{0, 1\}$	charging/discharging state of the storage unit during hour j

$x_j^{STOR} \in \mathbb{R}_+$	output power of the storage unit during hour j
$l_j^{STOR} \in \mathbb{R}_+$	energy capacity of the storage unit during hour j
$p_j^{STOR} \in \mathbb{R}_+$	energy inflows to the storage unit (or hydro reservoir) during hour j
$x_j^{DSM} \in \mathbb{R}_+$	energy 'put back' from demand side management during hour j
$l_j^{DSM} \in \mathbb{R}_+$	total energy withheld in demand side management during hour j
$p_j^{DSM} \in \mathbb{R}_+$	energy withheld in demand side management during hour j
$r_j^{PRI\_U}$	upwards primary reserves in hour j
$r_j^{PRI\_D}$	downwards primary reserves in hour j
$r_j^{SEC\_U}$	upwards secondary reserves in hour j
$r_j^{SEC\_D}$	downwards secondary reserves in hour j
r_j^{TER}	tertiary reserves in hour j
n_j	non-served energy in hour j

2.2.3 Parameters

Table 2.4: IMRES' Parameters

D_j	electricity demand in hour j [GWh]
CF_j^{WIND}	capacity factor of wind power during hour j [%]
CF_j^{SOLAR}	capacity factor of solar power during hour j [%]
C_i^{FOM}	fixed cost for operations and maintenance for unit i [M$/year]
C_i^{VOM}	variable cost for operations and maintenance for unit i [k$/MWh]
C_i^{FCAP}	annualized fixed cost for unit i [M$/year]
C_i^{FUEL}	fuel cost per unit power output from unit i
C_i^{STUP}	start-up cost for unit i
$VOLL$	value of lost load [k$/MWh]
Y_i	building state of renewable plant i: built ($Y_i = 1$) or not built ($Y_i = 0$)
E^{MAX}	limit on carbon emissions [Mtons]
E_i	carbon emissions per unit power output from power plant i [tons/GWh]
π^{CO_2}	price of carbon emissions [$/tn]

P_i^{MAX}	maximum power output for unit i [GW]
P_i^{MIN}	minimum stable power output for unit i [GW]
P^{IN}	maximum input power of storage unit (pumping capacity, in hydro) [GW]
P^{OUT}	maximum output power of storage unit [GW]
S_j	energy spilt in the storage unit during hour j
L^{MAX}	energy capacity of the aggregated storage unit [GWh]
L^{MIN}	minimum energy level of the aggregated storage unit [GWh]
I_j	hydro inflows during hour j [GWh]
ϵ	efficiency of the pumping unit in the reservoir [p.u.]
H	max time to 'put back' energy withheld in DSM
P^{DSM}	maximum DSM capacity in one hour
R_i^U	maximum up-ramping capability for unit i [GW/hr]
R_i^D	maximum down-ramping capability for unit i [GW/hr]
M_i^U	minimum up time for unit i
M_i^D	minimum down time for unit i
K^U	up reserves constant
K^D	down reserves constant
α	demand component for up primary reserves
β	demand component for down primary reserves
γ	wind output component for up secondary reserves
δ	solar output component for up primary reserves
ξ	wind output component for down secondary reserves
θ	solar output component for down secondary reserves
ζ	wind component for tertiary reserves
o	solar component for tertiary reserves

2.3 Model Formulation

$$\min_{\substack{x,y,z\\u,n,f,p}} \quad \sum_{i\in\mathcal{I}}(C_i^{FCAP}+C_i^{FOM})y_i+\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}}((C_i^{VOM}+C_i^{FUEL}+\pi^{CO_2}E_i)x_{ij}+C_i^{STUP}z_{ij}) + \sum_{j\in\mathcal{J}} VOLL\ n_j \quad (2.1)$$

s.t.

$$x_{ij} \leqslant P_i^{MAX}\ y_i \qquad \forall i\in\mathcal{I},\ \forall j\in\mathcal{J} \quad (2.2)$$

$$\sum_{i\in\mathcal{I}} x_{ij}+x_j^{STOR}+p_j^{DSM}+n_j = D_j+p_j^{STOR}+x_j^{DSM} \qquad \forall j\in\mathcal{J} \quad (2.3)$$

$$u_{ij}-u_{ij-1} = z_{ij}-v_{ij} \qquad \forall i\in\mathcal{T},\ \forall j\in\mathcal{J}\backslash\{1\} \quad (2.4)$$

$$w_{ij} = x_{ij}-u_{ij}\ P_i^{MIN} \qquad \forall i\in\mathcal{T},\ \forall j\in\mathcal{J} \quad (2.5)$$

$$w_{ij}-w_{ij-1} \leqslant R_i^U \qquad \forall i\in\mathcal{T},\ \forall j\in\mathcal{J}\backslash\{1\} \quad (2.6)$$

$$w_{ij-1}-w_{ij} \leqslant R_i^D \qquad \forall i\in\mathcal{T},\ \forall j\in\mathcal{J}\backslash\{1\} \quad (2.7)$$

$$w_{ij} \leqslant u_{ij}(P_i^{MAX}-P_i^{MIN}) \qquad \forall i\in\mathcal{T},\ \forall j\in\mathcal{J} \quad (2.8)$$

$$u_{ij} \geqslant \sum_{j'>j-M_i^U}^{j} z_{tj'} \qquad \forall i\in\mathcal{T},\ \forall j\in\mathcal{J} \quad (2.9)$$

$$1-u_{ij} \geqslant \sum_{j'>j-M_i^D}^{j} v_{tj'} \qquad \forall i\in\mathcal{T},\ \forall j\in\mathcal{J} \quad (2.10)$$

$$u_{ij} \geqslant y_i \qquad \forall i\in\mathcal{N},\forall j\in\mathcal{J} \quad (2.11)$$

$$y_i = Y_i \qquad \forall i\in\mathcal{W}\cup\mathcal{S} \quad (2.12)$$

$$\sum_{i\in\mathcal{W}} x_{ij} \leqslant \sum_{i\in\mathcal{W}} y_i\ P_i^{MAX}\ CF_j^{WIND} \qquad \forall j\in\mathcal{J} \quad (2.13)$$

$$\sum_{i\in\mathcal{S}} x_{ij} \leqslant \sum_{i\in\mathcal{S}} y_i\ P_i^{MAX}\ CF_j^{SOLAR} \qquad \forall j\in\mathcal{J} \quad (2.14)$$

$$\sum_{i\in\mathcal{I}}(E_i\sum_{j\in\mathcal{J}} x_{ij}) \leqslant E^{MAX} \quad (2.15)$$

$$l_j+S_j = l_{j-1}-x_j^{STOR}+I_j+\epsilon\ p_j \qquad \forall j\in\mathcal{J} \quad (2.16)$$

$$x_j^{STOR} \leqslant f_j P^{OUT} \qquad \forall j \in \mathcal{J} \qquad (2.17)$$

$$p_j \leqslant (1 - f_j) P^{IN} \qquad \forall j \in \mathcal{J} \qquad (2.18)$$

$$l_j \leqslant L^{MAX} \qquad \forall j \in \mathcal{J} \qquad (2.19)$$

$$l_j \geqslant L^{MIN} \qquad \forall j \in \mathcal{J} \qquad (2.20)$$

$$\sum_{j'>j}^{j+H} x_{j'}^{DSM} \geqslant p_j^{DSM} \qquad \forall j \in \mathcal{J} \qquad (2.21)$$

$$l_j^{DSM} = l_{j-1}^{DSM} - x_j^{DSM} + p_j^{DSM} \qquad \forall j \in \mathcal{J} \qquad (2.22)$$

$$l_j^{DSM} \leqslant P^{DSM} H \qquad \forall j \in \mathcal{J} \qquad (2.23)$$

$$x_j^{DSM} \leqslant P^{DSM} \qquad \forall j \in \mathcal{J} \qquad (2.24)$$

$$p_j^{DSM} \leqslant P^{DSM} \qquad \forall j \in \mathcal{J} \qquad (2.25)$$

$$r_j^{PRI\_U} \geqslant \max_{i \in \mathcal{I}} P_i^{MAX} \qquad \forall j \in \mathcal{J} \qquad (2.26)$$

$$r_j^{SEC\_U} \geqslant K^U [\alpha D_j^2 + \gamma (CF_j^{WIND} \sum_{i \in \mathcal{W}} Y_i \; P_i^{MAX})^2 + \\ + \delta (CF_j^{SOLAR} \sum_{i \in \mathcal{S}} Y_i \; P_i^{MAX})^2]^{1/2} \qquad \forall j \in \mathcal{J} \qquad (2.27)$$

$$r_j^{SEC\_D} \geqslant K^D [\beta D_j^2 + \xi (CF_j^{WIND} \sum_{i \in \mathcal{W}} Y_i \; P_i^{MAX})^2 + \\ + \theta (CF_j^{SOLAR} \sum_{i \in \mathcal{S}} Y_i \; P_i^{MAX})^2]^{1/2} \qquad \forall j \in \mathcal{J} \qquad (2.28)$$

$$r_j^{TER} \geqslant \zeta \sum_{i \in \mathcal{W}} Y_i \; P_i^{MAX} + o \sum_{i \in \mathcal{S}} Y_i \; P_i^{MAX} \qquad \forall j \in J \qquad (2.29)$$

$$\sum_{i \in \mathcal{I}} (u_{ij} \; P_i^{MAX} - x_{ij}) + f_j P^{OUT} - x_j^{STOR} + p_j + \\ + P^{DSM} + x_j^{DSM} - p_j^{DSM} \geqslant r_j^{PRI\_U} + r_j^{SEC\_U} \qquad \forall j \in \mathcal{J} \qquad (2.30)$$

$$\sum_{i \in \mathcal{I}} (x_{ij} - u_{i,j} \; P_i^{MIN}) + (1 - f_j) P^{OUT} + x_j^{STOR} - p_j \geqslant \\ r_j^{SEC\_D} \qquad \forall j \in \mathcal{J} \qquad (2.31)$$

$$\sum_{i \in \mathcal{I}} (y_i - u_{ij}) P_i^{MAX}) \geqslant r_j^{TER} \qquad \forall j \in \mathcal{J} \qquad (2.32)$$

$$x_{ij} \geqslant 0 \qquad \forall i \in \mathcal{I}, \; \forall j \in \mathcal{J} \qquad (2.33)$$

$$v_{ij} \geqslant 0 \qquad \forall i \in \mathcal{I},\ \forall j \in \mathcal{J} \qquad (2.34)$$

$$w_{i,j} \geqslant 0 \qquad \forall i \in \mathcal{I},\ \forall j \in \mathcal{J} \qquad (2.35)$$

$$l_j^{STOR} \geqslant 0 \qquad \forall j \in \mathcal{J} \qquad (2.36)$$

$$l_j^{DSM} \geqslant 0 \qquad \forall j \in \mathcal{J} \qquad (2.37)$$

$$p_j^{STOR} \geqslant 0 \qquad \forall j \in \mathcal{J} \qquad (2.38)$$

$$p_j^{DSM} \geqslant 0 \qquad \forall j \in \mathcal{J} \qquad (2.39)$$

$$n_j \geqslant 0 \qquad \forall j \in \mathcal{J} \qquad (2.40)$$

$$y_i \in \{0,1\} \qquad \forall i \in \mathcal{I} \qquad (2.41)$$

$$f_i \in \{0,1\} \qquad \forall i \in \mathcal{I} \qquad (2.42)$$

$$u_{ij} \in \{0,1\} \qquad \forall i \in \mathcal{I},\ \forall j \in \mathcal{J} \qquad (2.43)$$

$$z_{ij} \in \{0,1\} \qquad \forall i \in \mathcal{I},\ \forall j \in \mathcal{J} \qquad (2.44)$$

2.4 Description of the Model

2.4.1 Indices and Sets

IMRES uses two main indices: $i \in \mathcal{I}$ and $j \in \mathcal{J}$. $\mathcal{I}$ denotes the set of individual generating units that can be built, containing all the units from the different technologies considered: nuclear, coal, combined cycle gas turbines (CCGTs), open cycle gas turbines (OCGTs), wind and solar photovoltaic. In addition, $\mathcal{W} \subset \mathcal{I}$ denotes the subset of wind units; $\mathcal{S} \subset \mathcal{I}$, denotes the subset of solar units; $\mathcal{T} \subset \mathcal{I}$, denotes the subset of thermal power units (nuclear, coal, CCGTs and OCGTs); and $\mathcal{G} \subset \mathcal{I}$, denotes the subset of gas-fired power plants (CCGTs and OCGTs). $\mathcal{J}$ denotes the set of hours in a year or, alternatively, the total number of hours contained in the weeks sampled, used in the unit commitment component of the model.

Building decisions are modeled with binary variables $y_i \in \{0,1\}$, where $\boldsymbol{y}^*$ is a vector representing the optimal building state of the power plants in the solution space; unit commitment decisions are denoted by $u_{i,j} \in \{0,1\}$; start-up decisions are denoted by $z_{ij} \in \{0,1\}$; shut-down decisions are denoted by $v_{ij} \in \mathbb{R}_+$; power output decisions are denoted by

$x_{ij} \in \mathbb{R}_+$; and non-served energy is denoted by $n_j \in \mathbb{R}_+$. An additional variable $w_{ij} \in \mathbb{R}_+$ (where $w_{ij} = x_{ij} - P_i^{MIN} u_{ij}$), has been introduced to decompose the total output of each power plant into its minimum stable level and the incremental generation above that minimum output to facilitate the formulation of ramping rate constraints.

2.4.2 Objective Function

The objective function in IMRES (2.45) minimizes the total cost of the system, which is the sum of fixed costs (annualized capital cost plus fixed O&M: $C_i^{FCAP} + C_i^{FOM}$), variable costs[1] (fuel cost plus variable O&M: $C_i^{FUEL} + C_i^{VOM}$), start-up costs (C_i^{STUP}) and the value of lost load (C^{VOLL}). IMRES can formally be divided into two components: 1) a component where individual plant building decisions are made; and 2) a component including startup, commitment and energy dispatch decisions. Note however that these components are considered jointly and the optimization is performed over the whole problem at once.

$$\min_{\substack{x,y,z \\ u,n,f,p}} \quad \sum_{i\in\mathcal{I}}(C_i^{FCAP} + C_i^{FOM})y_i + \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}}((C_i^{VOM} + C_i^{FUEL} + \pi^{CO_2}E_i)x_{ij} + C_i^{STUP} z_{ij}) + \sum_{j\in\mathcal{J}} VOLL\ n_j \qquad (2.45)$$

2.4.3 Value of Lost Load

In the same way as in other capacity expansion models, the value of lost load (VOLL) used in IMRES affects critically the total non-served energy in the system and building decisions of peaking units. The choice of VOLL can be made to address various criteria and there are several methods to determine its value.

To give an example, a typical rule of thumb for establishing reliability criteria in many power systems is that there can only be at most one day (24 hrs of duration) with non-served energy over a time span of ten years. If we calculate the per year ratio of this limitation, we

[1]The present formulation uses an affine variable cost function that could be replaced in future implementations by a piecewise linear cost function to increase the accuracy of the representation of the plants' cost structure.

have a maximum average of 2.4 hours per year with non-served energy. In a classic capacity expansion model with screening curves (Turvey and Anderson, 1977), non-served energy can be introduced as an additional technology with fixed cost equal to zero and variable cost equal to the VOLL. The intersection between the cost function of non-served energy and the cost function of the peaking technology determines the number of hours in a year for which it is cheaper to curtail demand rather than supply the full peak. If the maximum number of hours with non-served energy is fixed by the reliability criteria of choice (for the criteria just discussed, 2.4 hours), and we assume that the peaking plant does not have any additional income, we can derive an analytical expression for the VOLL using this number:

$$VOLL = \frac{C_i^{FCAP} + C_i^{FOM}}{2.4} + C_i^{VOM} + C_i^{FUEL}, \quad i \in \mathcal{G} \qquad [\$/MWh] \qquad (2.46)$$

Conventional values of lost load range between 1,000 and 10,000 $/MWh. However, the method just described typically produces values much higher than 10,000 $/MWh. For instance, for a peaking technology with a fixed annual cost of 100,000 $/MW, the resulting value of lost load calculated with screening curves is above 40,000 $/MWh.

Alternatively, we can assume that some demand is sensitive to price, which prevents the price from going above some certain threshold lower than the values presented above. In reality, this is achieved through contracts with special customer groups that are willing to reduce their demand during peak hours, or with grid elements that can supply electricity on an ad-hoc basis. Elements within this category are emergency generators located in critical infrastructures and public facilities such as hospitals, government offices, etc, used for back-up power in case of blackouts, staying idle when the system is operating normally. These generators could potentially be used to deliver electricity when prices are high, without jeopardizing their back-up generator functionality. Typically, back-up generators are fueled with expensive diesel, and if they are used in the mode just described, the system VOLL would take a value close to the variable cost of these generators, but high enough to offer an incentive for these generators to supply electricity to the system. The analyses in this thesis assume a VOLL$\sim$ 500 $/MWh, which is close to the variable cost of these plants.

2.4.4 Accounting for Unit Commitment

The main two decision components in IMRES (building and operating) are linked with a coupling constraint imposing the condition that only units that have been built can generate (3.17).

$$x_{ij} \leqslant P_i^{MAX}\, y_i \qquad \forall i \in \mathcal{I},\ \forall j \in \mathcal{J} \tag{3.17}$$

Additionally, IMRES is subject to the classic constraints included in a unit commitment model: demand balance constraints; constraints on the commitment state; constraints on the minimum and maximum output of the plants; constraints on the ramping rates; and constraints on the minimum up and down times . The demand balance constraint (3.25) establishes for all time periods the equilibrium between the load in the system and the total power generated, including the option of having non-served energy.

$$\sum_{i \in \mathcal{I}} x_{ij} + n_j = D_j \qquad \forall j \in \mathcal{J} \tag{2.47}$$

If we include the possibility of having storage and demand side management (DSM) in the system (see sections 2.4.6 and 2.4.7), the demand balance equation becomes:

$$\sum_{i \in \mathcal{I}} x_{ij} + x_j^{STOR} + p_j^{DSM} + n_j = D_j + p_j^{STOR} + x_j^{DSM} \qquad \forall j \in \mathcal{J} \tag{2.3}$$

State constraints (3.5) link commitment states (u_{ij}) with start-up and shut-down decisions (z_{ij} and v_{ij}, respectively). Note that even if $v_{i,j}$ has been defined in the positive real domain, it will only adopt binary values as the commitment states and start-up decisions are all binary.

$$u_{ij} - u_{ij-1} = z_{ij} - v_{ij} \qquad \forall i \in \mathcal{T},\ \forall j \in \mathcal{J} \backslash \{1\} \tag{3.5}$$

Constraints on ramping rates (3.7-3.8) account for the physical limitations imposed by

power plants' thermal and mechanical inertias. These equations are constructed using a set of auxiliary variables ($w_{ij}, i \in \mathcal{T}, j \in \mathcal{J}$) (3.6) to prevent ramping constraints from acting when off-line power plants start-up and jump from zero output ($x_{ij} = 0$) to the minimum output ($x_{ij} = P_i^{MIN}$), or when power plants shut down.

$$w_{ij} = x_{ij} - u_{ij}\, P_i^{MIN} \qquad \forall i \in \mathcal{T},\ \forall j \in \mathcal{J} \tag{3.6}$$

$$w_{ij} - w_{ij-1} \leqslant R_i^U \qquad \forall i \in \mathcal{T},\ \forall j \in \mathcal{J}\backslash\{1\} \tag{3.7}$$

$$w_{ij-1} - w_{ij} \leqslant R_i^D \qquad \forall i \in \mathcal{T},\ \forall j \in \mathcal{J}\backslash\{1\} \tag{3.8}$$

Constraints accounting for the minimum and maximum output limits of the units (3.9) can be thus defined in terms of the interval between each unit's minimum and maximum output levels:

$$w_{ij} \leqslant u_{ij} \left(P_i^{MAX} - P_i^{MIN}\right) \qquad \forall i \in \mathcal{T},\ \forall j \in \mathcal{J} \tag{3.9}$$

Constraints on the minimum up and down times are implemented following the formulation in (Hedman et al., 2009). In this formulation M_i^U and M_i^D represent the minimum time that a power plant has to remain on or off after a start-up or shut-down respectively, and $j' \in \mathcal{J}$ is an auxiliary index for the hours in the time series:

$$u_{ij} \geqslant \sum_{j' > j - M_i^U}^{j} z_{tj'} \qquad \forall i \in \mathcal{T},\ \forall j \in \mathcal{J} \tag{3.10}$$

$$1 - u_{ij} \geqslant \sum_{j' > j - M_i^D}^{j} v_{tj'} \qquad \forall i \in \mathcal{T},\ \forall j \in \mathcal{J} \tag{3.11}$$

Lastly, IMRES allows the option of including a restriction on nuclear plants cycling that prevents nuclear plants from shutting down or starting up for reasons other than for maintenance, as is the case in many power systems. If this constraint is added to the

formulation, all nuclear power plants built have to be permanently on-line (3.12).

$$u_{ij} \geqslant y_i \qquad \forall i \in \mathcal{N},\ \forall j \in \mathcal{J} \tag{3.12}$$

2.4.5 Accounting for Renewables and Carbon Emissions

Wind and solar PV plants ($Y_i,\ i \in \mathcal{W} \cup \mathcal{S}$) are considered as inputs to the model, and their energy outputs are functions of each technology's capacity factor (CF). Capacity factors reflect the availability of wind or solar resources for a specific hour at a certain location. Hence, for a particular hour of the year, the output from the total wind or solar power installed in the system is determined by the product of each technology's total capacity installed and its respective CF (3.16-2.49):

$$y_i = Y_i \qquad \forall i \in \mathcal{W} \cup \mathcal{S} \tag{2.12}$$

$$\sum_{i \in \mathcal{W}} x_{ij} = \sum_{i \in \mathcal{W}} y_i\ P_i^{MAX}\ CF_j^{WIND} \qquad \forall j \in \mathcal{J} \tag{2.48}$$

$$\sum_{i \in \mathcal{S}} x_{ij} = \sum_{i \in \mathcal{S}} y_i\ P_i^{MAX}\ CF_j^{SOLAR} \qquad \forall j \in \mathcal{J} \tag{2.49}$$

–$P_i^{MAX}, i \in \mathcal{W}$ and $P_i^{MAX}, i \in \mathcal{S}$ are the maximum capacities of a unitary wind and solar farm respectively (in this research, equal to 1GW). Note that in IMRES the size of renewable power plants –wind and solar PV– does not affect the outcomes of the model, as long as the sum of the capacity of individual wind and solar plants in the system adds up separately to the total wind and solar capacity in place in the scenario analyzed.

The model also allows the possibility of introducing curtailment as another degree of freedom to guarantee that generation exactly meets hourly demand. For both wind and solar power, this feature is implemented through substituting the equality constraints (3.16-

2.49) by inequality constraints (3.13-3.14):

$$\sum_{i \in \mathcal{W}} x_{ij} \leqslant \sum_{i \in \mathcal{W}} y_i \; P_i^{MAX} \; CF_j^{WIND} \qquad \forall j \in \mathcal{J} \tag{3.13}$$

$$\sum_{i \in \mathcal{S}} x_{ij} \leqslant \sum_{i \in \mathcal{S}} y_i \; P_i^{MAX} \; CF_j^{SOLAR} \qquad \forall j \in \mathcal{J} \tag{3.14}$$

Additionally, each unit has a parameter E_i reflecting an estimation of its specific carbon emissions in gCO_2/kWh associated with producing electricity. The effect of these carbon emissions is modeled by either introducing a carbon price π^{CO_2} [$/tn$CO_2$] in the objective function that affects the total variable cost of each technology, or by introducing a cap on total system emissions E^{MAX}, implemented as an additional constraint (3.15).

$$\sum_{i \in \mathcal{I}} \left(E_i \sum_{j \in \mathcal{J}} x_{ij} \right) \leqslant E^{MAX} \tag{3.15}$$

Note that if the latter option is used, the dual variable of the emissions constraints can be interpreted as the carbon price that plants should be charged in a carbon price scenario to achieve the same emissions target of E^{MAX} equivalent tons of CO_2 .

2.4.6 Accounting for Hydro and Storage Capacity

IMRES aggregates hydro storage and other forms of storage capacity (mostly electro-chemical and chemical) in a single storage unit with the ability to increase and release the energy stored. The formulation used in this model is similar to that used by Cerisola et al. (Cerisola et al., 2009), where a constant energy-flow ratio for each hydro unit is considered and the storage level is expressed in terms of stored energy in MWh. The equation representing the storage level of the reservoir is as follows:

$$l_j^{STOR} + S_j = l_{j-1}^{STOR} - x_j^{STOR} + I_j + \epsilon \; p_j^{STOR} \qquad \forall j \in \mathcal{J} \tag{2.16}$$

–where $l_j^{STOR} \in \mathbb{R}_+$ represents the energy level of the storage unit at hour j; $p_j^{STOR} \in \mathbb{R}_+$

represents the pumped energy during hour j; ϵ is the efficiency of the pumping unit; S_j is the energy spilled in hour j; and I_j are the energy inflows (water inflows from precipitation, in the case of hydro storage) during hour j.

Additionally, storage units are subject to constraints on the maximum and minimum level that the storage unit can reach, and the maximum speed of charging and discharging the unit. These constraints are respectively:

$$l_j^{STOR} \leqslant L^{MAX} \qquad \forall j \in \mathcal{J} \tag{2.17}$$

$$l_j^{STOR} \geqslant L^{MIN} \qquad \forall j \in \mathcal{J} \tag{2.18}$$

$$p_j^{STOR} \leqslant (1 - f_j)\, P^{IN} \qquad \forall j \in \mathcal{J} \tag{2.19}$$

$$x_j^{STOR} \leqslant f_j P^{OUT} \qquad \forall j \in \mathcal{J} \tag{2.20}$$

–where L^{MAX} and L^{MIN} denote the maximum and minimum storage level of the aggregated unit; P^{IN} and P^{OUT} are the maximum speed to discharge and charge the unit, with $P^{OUT} \leqslant L^{MIN}$; and the variable $f_j \in \{0, 1\}$ denotes the charging or discharging state of the storage unit ($f_j = 0$ indicates that the storage unit is charging, and $f_j = 1$ indicates that the storage unit is discharging).

An alternative to aggregating all storage units into a single unit is to treat them separately, splitting them into different units. Each unit can be distinguished from the other through different values of storage capacity, maximum power delivery or roundtrip efficiency. However, introducing many storage units in the model will have a direct impact on the model's computational time, as the number of decision variables increase.

2.4.7 Accounting for Demand Side Management

Demand side management (DSM) is a strategy designed to withhold a fraction of the demand until a later time, when it is cheaper for the system to supply that energy. DSM programs try to achieve net load peak shaving and provide reserves, which would reduce fuel costs and defer investments.

In IMRES, DSM has been implemented as a capability to shift part of the demand at a given hour throughout the following H hours. The implementation of DSM is somewhat similar to the implementation of storage as the amount of demand that remains to be supplied is recorded at a 'virtual storage' unit. This virtual storage can only store a maximum of $P^{DSM}H$ gigawatt-hours of energy, where P^{DSM} is the maximum energy that can be withheld in one hour. Accordingly, the equation modeling the behavior of DSM is:

$$\sum_{j'>j}^{j+H} x_{j'}^{DSM} \geqslant p_j^{DSM} \qquad \forall j \in \mathcal{J} \qquad (2.21)$$

$$l_j^{DSM} = l_{j-1}^{DSM} - x_j^{DSM} + p_j^{DSM} \qquad \forall j \in \mathcal{J} \qquad (2.22)$$

$$l_j^{DSM} \leqslant P^{DSM}H \qquad \forall j \in \mathcal{J} \qquad (2.23)$$

$$x_j^{DSM} \leqslant P^{DSM} \qquad \forall j \in \mathcal{J} \qquad (2.24)$$

$$p_j^{DSM} \leqslant P^{DSM} \qquad \forall j \in \mathcal{J} \qquad (2.25)$$

–where $p_j^{DSM} \in \mathbb{R}_+$ denotes the energy withheld during hour j; $l_j^{DSM} \in \mathbb{R}_+$ denotes the total energy withheld or, alternatively, the energy level of the virtual storage at hour j; P^{DSM} denotes the demand shifting capacity; and x_j^{DSM} denotes the demand being put back at hour j.

2.4.8 Reserves Constraints

IMRES uses the general reserves classification proposed in Milligan et al. (Milligan et al., 2010), but with the European naming convention (primary, secondary and tertiary reserves, as in Figure 1-2). Primary up reserves, are denoted $r_j^{PRI\_U}$, and are equal to the capacity of the largest generating unit. The magnitude of secondary up reserves, $r_j^{SEC\_U}$, and of secondary down reserves, $r_j^{SEC\_D}$, is proportional to the euclidean norm of three uncertainty factors, depending on the demand level, and the wind and solar PV outputs[2] (Milligan et al., 2010). Finally, tertiary reserves, are denoted r_j^{TER}, are proportional to the wind and solar

[2]This assumes that the forecasting errors associated with these three factors (the sources of uncertainty) are independent from each other.

PV capacity installed.

$$r_j^{PRI\_U} \geqslant \max_{i \in \mathcal{I}} P_i^{MAX} \qquad \forall j \in \mathcal{J} \tag{2.26}$$

$$r_j^{SEC\_U} \geqslant K^U \left[\alpha D_j^2 + \gamma \left(CF_j^{WIND} \sum_{i \in \mathcal{W}} Y_i \; P_i^{MAX} \right)^2 + \delta \left(CF_j^{SOLAR} \sum_{i \in \mathcal{S}} Y_i \; P_i^{MAX} \right)^2 \right]^{1/2} \qquad \forall j \in \mathcal{J} \tag{2.27}$$

$$r_j^{SEC\_D} \geqslant K^D \left[\beta D_j^2 + \xi \left(CF_j^{WIND} \sum_{i \in \mathcal{W}} Y_i \; P_i^{MAX} \right)^2 + \theta \left(CF_j^{SOLAR} \sum_{i \in \mathcal{S}} Y_i \; P_i^{MAX} \right)^2 \right]^{1/2} \qquad \forall j \in \mathcal{J} \tag{2.28}$$

$$r_j^{TER} \geqslant \zeta \sum_{i \in \mathcal{W}} Y_i \; P_i^{MAX} + o \sum_{i \in \mathcal{S}} Y_i \; P_i^{MAX} \qquad \forall j \in \mathcal{J} \tag{2.29}$$

Operating reserves are modeled requiring the system to have a certain amount of spinning and non-spinning reserves ready to be deployed at every hour. These reserves can be provided by committed thermal power plants, storage units and DSM. Accordingly, spinning reserves are formulated accounting for the capacity margin offered by these on-line elements (2.30-2.31).

$$\sum_{i \in \mathcal{I}} \left(u_{ij} \; P_i^{MAX} - x_{ij} \right) + f_j P^{OUT} - x_j^{STOR} + p_j^{STOR} + P^{DSM} + \\ + x_j^{DSM} - p_j^{DSM} \geqslant r_j^{PRI\_U} + r_j^{SEC\_U} \qquad \forall j \in \mathcal{J} \tag{2.30}$$

$$\sum_{i \in \mathcal{I}} \left(x_{ij} - u_{ij} \; P_i^{MIN} \right) + (1 - f_j) \, P^{OUT} + x_j^{STOR} - p_j^{STOR} + P^{DSM} - \\ - x_j^{DSM} + p_j^{DSM} \geqslant r_j^{SEC\_D} \qquad \forall j \in \mathcal{J} \tag{2.31}$$

Additional conditions should be added to these two equations in order to take into account the behavior of the reservoir when the storage is at minimum or maximum capacity: in Equation (2.30), if $l_j^{STOR} - L^{MIN} < P^{OUT}$, then P^{OUT} should be replaced by $l_j^{STOR} - L^{MIN}$; similarly, in Equation (2.31), if $L^{MAX} - l_j^{STOR} < P^{OUT}$, then P^{OUT} should be replaced by $L^{MAX} - l_j^{STOR}$.

Finally, non-spinning reserves are modeled accounting for the off-line capacity in the system that can be turned on to replace secondary reserves (2.32).

$$\sum_{i \in \mathcal{I}} (y_i - u_{i,j})\, P_i^{MAX} \geqslant r_j^{TER} \qquad \forall j \in \mathcal{J} \qquad (2.32)$$

2.5 Approximating the Net Load Duration Curve

IMRES is formulated as a high-dimension *binary mixed integer linear program* (0-1 MILP), in which unit commitment states and investment decisions are represented by binary variables. The dimensionality of a formulation like the one proposed, combining investment decisions at the individual plant level and a unit commitment spanning one full year, is on the order of $8\ 760 \cdot \dim(\boldsymbol{y})$, where $\dim(\boldsymbol{y})$ is the number of power plants considered by the model. For realistically sized system (with a peak load on the order of tens of gigawatts), $\dim(\boldsymbol{y})$ will typically be in the order of hundreds, which means that the computational burden of solving the model is prohibitively large for current computers and solvers such as CPLEX. Decomposition techniques such as Benders' cannot be used with this formulation because the binary variables in the commitment component (the subproblem in a decomposed formulation) would impede the closing of the duality gap (Floudas, 1995). Therefore, simplifications are needed to reduce the dimensionality of the formulation.

The characteristic variability of power systems with high shares of renewables can be approximated using IMRES by introducing a four-week approximation of a one-year net load series to reflect the joint variability of demand and renewable energy output (note that longer time series could also be used in this process to account for a larger variability spectrum). The week selection heuristic method is based on choosing the four weeks that

most closely approximate the full annual net load duration curve, using least-squares. A detailed description of the week selection heuristic method used in IMRES, its performance and its comparative advantage against other selection methods is presented below.

2.5.1 Week Selection Methodology

The methodology presented in this section is based on selecting a fixed number of sample weeks from a one-year-horizon net load series, to develop useful approximations to be used within planning tools for systems with a very high penetration of renewables. To this end, we need to assume the hourly renewable generation for a certain future penetration level as an input to the model. The method focuses on net load because it represents the combined variability of demand and renewables that must be balanced with dispatchable generation (mostly thermal and hydro units). Therefore, by jointly considering demand and renewable generation, we also account for the correlation between them.

It is well-known that there is a smoothing effect on the variability of renewable output with a larger geographical dispersion of the resources. However, for sufficiently large levels of dispersion and penetration, this smoothing effect reaches a saturation point above which there is minimal further reduction in variability (Hasche, 2010). This heuristic method assumes that for the systems studied such a saturation point has been reached and that it is safe to scale historical data to project hourly renewable output for higher levels of penetration. Therefore, for a given penetration level of wind and solar PV, the hourly variable output is calculated as the product of each variable technology's historical capacity factor (CF) and the total capacity deployed of that technology. If we choose the maximum output parameter of renewable plants to be equal to 1GW ($P_i^{MAX} = 1\text{GW} \quad \forall i \in \mathcal{W} \cup \mathcal{S}$), then the gigawatts of wind and solar capacity installed in the system can be expressed as $\sum_{i \in \mathcal{W}} y_i$ and $\sum_{i \in \mathcal{S}} y_i$ respectively. Thus, the *net load duration curve* ($NLDC$) can be obtained by sorting the difference between the demand and the renewable output in decreasing order:

$$
\begin{aligned}
NLDC_j &= D_j - CF_j^{WIND} \sum_{i \in \mathcal{W}} y_i - CF_j^{SOLAR} \sum_{i \in \mathcal{S}} y_i \\
\text{s.t.:} \quad & NLDC_{j-1} - NLDC_j \geqslant 0 \qquad \forall j \in \mathcal{J}
\end{aligned}
\tag{2.50}
$$

Similarly, an *approximate net load duration curve* ($\widetilde{NLDC}$) can be obtained in three consecutive steps: 1) sample a given number of weeks from a full year of demand and renewable generation; 2) scale up the hours contained in the sample to one year; and 3) sort the series in decreasing order to form the approximate net load duration curve. The resulting ($\widetilde{NLDC}$) is a staircase approximation to the $NLDC$.

In this analysis, for a scenario given by its renewable penetration level, we choose the four weeks that minimize the least square error between the $NLDC$ and its approximation. The approximation (ν^*) is thus given by the solution to the following optimization problem:

$$
\nu^* = \arg\min_{\nu} \| NLDC - \widetilde{NLDC}^{\nu} \|^2 \tag{2.51}
$$

where $\nu \in \mathbb{Z}^n$ is the set of indices of the n weeks selected; $\nu^* \in \mathbb{Z}^n$ is the set of indices of the optimal week combination; and $NLDC, \widetilde{NLDC}^{\nu} \in \mathbb{R}^{8,736}$.

This method guarantees that the weeks selected best represent the energy below the net load duration curve, but it does not explicitly account for the chronological variability of the net load. Therefore, the method cannot guarantee that the four weeks selected provide a good representation of the chronological order of the net load in one year. Still, as it will be shown below, sensitivity analyses conducted on the number of weeks selected to construct the approximation $\widetilde{NLDC}$ and a thorough comparison of the commitment results produced by the approximation and the full year net load series show that, for the 4-week approximation ($\dim \nu = 4$), the errors are small both in terms of energy and in terms of commitment results (less than 3% in terms of capacity factor and energy contribution, and less than 10% of total startups).

2.5.2 Scaling the Sample

We can assess the quality of the approximation by comparing each of the individual hours in the $NLDC$ with its corresponding hour in the $\widetilde{NLDC}$. In order to establish this comparison, we need to have the same number of hours in the approximation as in the $NLDC$. To this end, each hour contained in the weeks sampled has to be repeated an integer number of times (Θ) to complete the total number of hours in a year. For simplicity, although one year is composed of 52.14 weeks, we will assume that a year is formed by 52 weeks to preserve the integrality of the total number of weeks selected. For example, if a sample of four weeks is selected, the factor by which we have to repeat each hour within the group of weeks selected to generate a one-year series is $\Theta = 52/4 = 13$ (each hour from the 4 weeks appears 13 times in the annual approximation) (Figure 2-1).

Notice that if the same factor is to be applied to all the weeks selected, the possible number of weeks is limited to the divisors of 52 (i.e., 1, 2, 4, 13, 26 and 52). Other different number of weeks could be selected, but applying different scaling factors to each of the weeks.

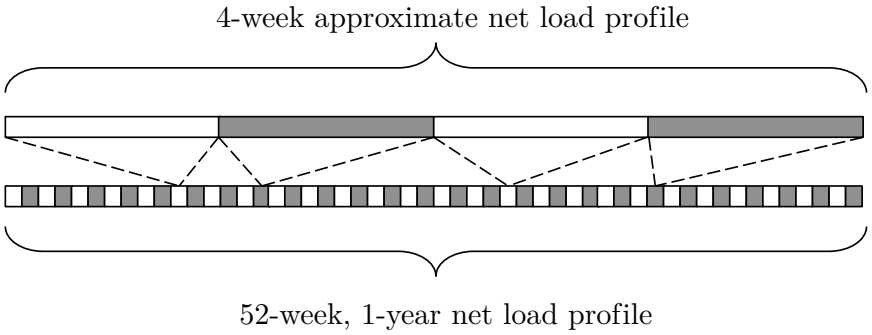


Figure 2-1: Four-week approximate net load constructed with 4 weeks selected from the total 52 weeks in one year. Each hour from the selected week is scaled by a factor $\Theta = 52/4 = 13$

2.5.3 Error Metrics

As mentioned above, the objective of the approximation is two-fold: first, to accurately approximate the yearly net energy demanded by the system; and second, to account for the operational detail of thermal units in the system. Thus, we need metrics that assess the

approximations performance from these two perspectives. The statistical metric for the error incurred with respect to the net energy is the root-mean-square error ($RMSE$) between the $NLDC$ and its approximation:

$$RMSE^{\nu} = \sqrt{\frac{\sum_{j=1}^{8,736} \left(NLDC_j - \widetilde{NLDC}_j^{\nu} \right)^2}{8,736}} \tag{2.52}$$

This metric can be normalized relative to the range of net load values in the series with the normalized-root-mean-square error ($NRMSE$):

$$NRMSE^{\nu} = \frac{RMSE^{\nu}}{NLDC_{max} - NLDC_{min}} \cdot 100 \quad [\%] \tag{2.53}$$

The sample of weeks is chosen to best represent the net load duration curve. However, we are primarily interested in how well the sample approximates the unit commitment decisions from a full year. To show the value of the approximation, we compare the results for energy contribution, capacity factor, and number of cycles performed per year from the approximation (scaled up to a year) with the results from running the UC for the full year. These three values give an indication of how thermal units are operated in the system. The difference in results between the two models is expressed in absolute values with their sign, and define the three error metrics for each technology $k \in \mathcal{K}$ as:

$$ECE^k = \frac{1}{|\mathcal{T}^k|} \left(\sum_{i \in \mathcal{T}^k} \widetilde{EC}_i - \sum_{i \in \mathcal{T}^k} EC_i \right) \qquad \text{[p.u.]} \tag{2.54}$$

$$CFE^k = \frac{1}{|\mathcal{T}^k|} \left(\sum_{i \in \mathcal{T}^k} \widetilde{CF}_i - \sum_{i \in \mathcal{T}^k} CF_i \right) \qquad \text{[p.u.]} \tag{2.55}$$

$$SUE^k = \frac{1}{|\mathcal{T}^k|} \left(\sum_{i \in \mathcal{T}^k} \widetilde{SU}_i - \sum_{i \in \mathcal{T}^k} SU_i \right) \qquad \text{[p.u.]} \tag{2.56}$$

where $\mathcal{T}^k \subset \mathcal{I}$ is the subset of indices corresponding to generating units of technology k; and $|\mathcal{T}^k|$ is the cardinality of the subset.

2.5.4 How Many Weeks to Select?

The selection algorithm consists of an exhaustive search throughout all the possible combinations of weeks to determine which combination yields the minimum RMSE. This enumeration process is implemented by a number of nested loops equal to the number of weeks to be selected. As a result, the computational time required to choose the optimal sample set grows proportionally with the number of possible combinations, which grows as the factorial of the number of sample weeks (Table 2.5).

Table 2.5: Number of Possible Combinations and Computing Time as a Function of the Number of Weeks Selected

Number of weeks Selected	Number of possible combinations	Computing time
1	$\binom{52}{1} = 52$	0.05 secs
2	$\binom{52}{2} = 1,326$	1.5 secs
4	$\binom{52}{4} = 270,725$	10 mins
5	$\binom{52}{5} = 2,598,960$	1 h 40 mins
8	$\binom{52}{8} = 752,538,150$	∼19 days
13	$\binom{52}{13} = 6.3501E+11$	∼ 46 years

Computing times from running a MATLAB algorithm having as many nested loops as weeks selected, on a commercial 2.4 GHz Intel Core 2 Duo machine with 4GB of memory.

Incrementing the sample size by one week requires an additional nested loop in the sampling algorithm to search through all possible week combinations. On a standard desktop, the optimization procedure can be reasonably performed for sample sizes of up to 5 weeks (5 nested loops), but beyond that it becomes prohibitive.

2.5.5 Selecting a Large Number of Weeks

As an alternative to optimal selection for larger sample sizes, there are heuristic approximations that allow increasing the sample size without increasing the number of nested loops. For instance, we could sample subsets of some certain number of adjoining weeks instead of individual weeks (e.g., two-week intervals). With this heuristic we can easily expand the

number of weeks selected but, for a given number of weeks, the weeks selected are not the ones representing the NLDC the best. Once these subsets have been selected, we can scale up the hours in the sample to produce a 52-week approximation by applying a different scaling factor to each week within the subset. For instance, a 16-week sample might consist of four subsets of four weeks, where the first week in the subset is scaled up by a factor of 6, the second week by a factor of 4, the third by a factor of 2, and the fourth by a factor of 1 $(4 \cdot (6 + 4 + 2 + 1) = 52)$:

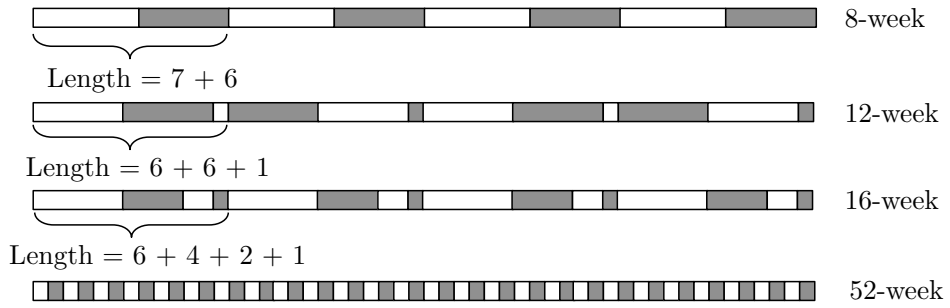


Figure 2-2: Four-week approximate net load constructed with 8, 12, 16 and 52 weeks selected from the total 52 weeks in one year.

This approach allows to increase the number of weeks sampled while avoiding a dramatic increase in computation time as the number of loops in the search process grow.

2.5.6 Number of Weeks and Net Energy Error

The demand and renewable generation (solar and wind power) hourly data used to test the proposed methodology corresponds to the data recorded in 2009 in the Electricity Reliability Council of Texas (ERCOT). In order to simulate a high penetration level of renewables, wind and solar outputs are scaled up assuming an expected capacity of 30GW of wind power and 10 GW of solar PV (Holttinen et al., 2010).

Figure 2-3 shows graphically the results of selecting one, two and four weeks to characterize the $NLDC$, illustrating how the RMSE in approximating the net load decrease with the number of weeks sampled.

It can be observed that while the approximations based on one and two weeks produce

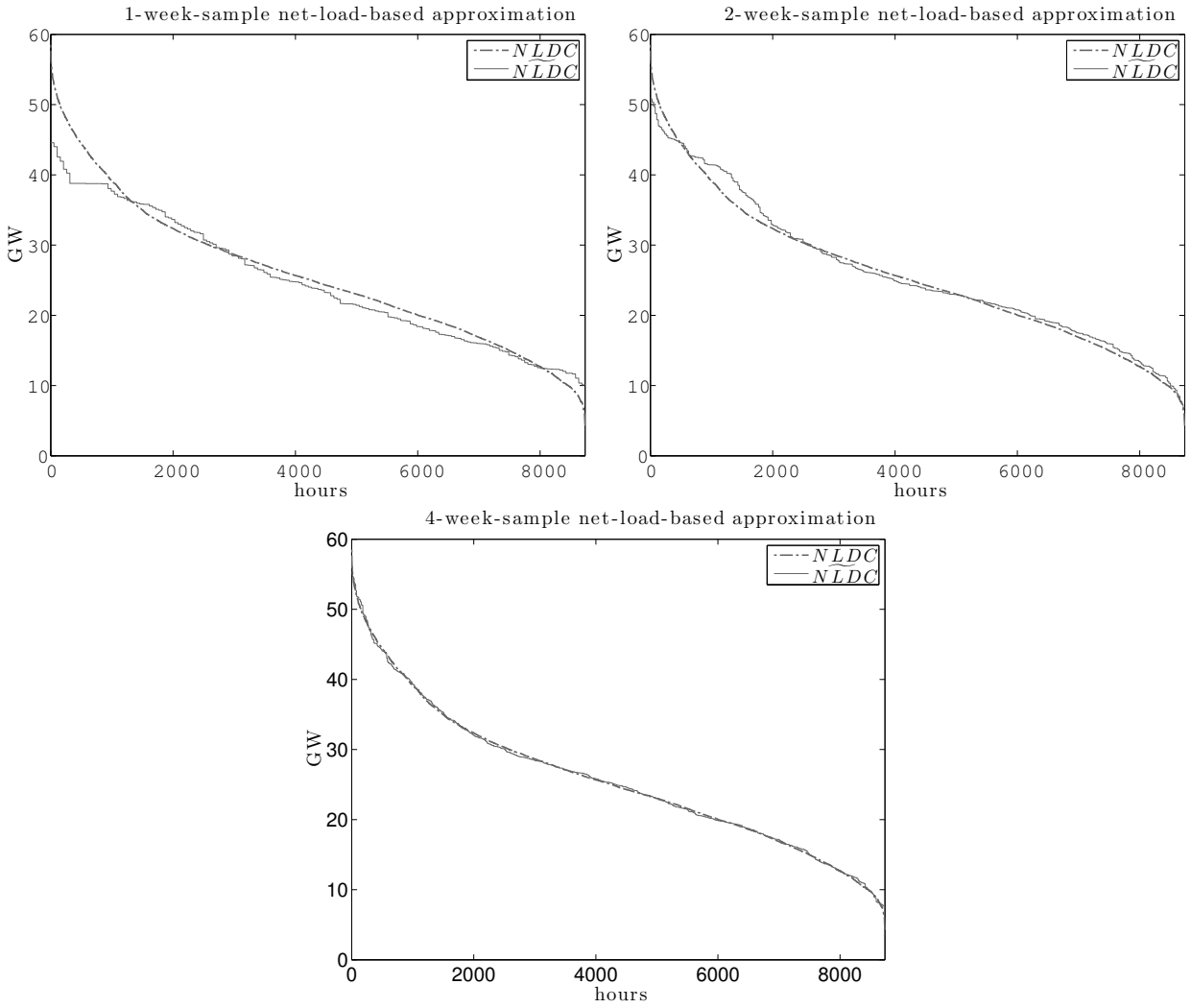


Figure 2-3: Graphical representation of the system's LDC, NLDC and their respective approximations, built with one, two and four weeks, selected to fit the NLDC.

an inaccurate representation of the NLDC with a RMSE over 2%, the approximation that uses four weeks closely matches the shape of the original NLDC. For sample sizes larger than four weeks, selected with the heuristic method that chooses from subsets of weeks (rather than individual weeks), errors are further reduced, although non-monotonically (Table 2.6).

Table 2.6: Net Energy Error Relative to the Number of Weeks Selected

Number of Weeks	Approximation ν*	$RMSE^{\nu}$	$NRMSE^{\nu}$ [%]	time [sec]
1	{4}	2.3	4.2	0.066
2	{9,24}	1.233	2.3	1.52
4	{17,31,37,46}	0.299	0.5	688
8	{3,4,15,16,20,21,31,32}	0.283	0.52	1,042
12	{12,13,14,34,35,36,38,39,40,50,51,52}	0.68	1.24	1,368
16	{3,4,5,6,14,15,16,17,31,32,33,34,38,39,40,41}	0.24	0.44	1,505

Summary of the error incurred by approximations constructed with an increasing number of weeks. The weeks are selected based on minimizing the error between the approximation and the NLDC. The column "Approximation" shows the weeks selected; "$RMSE^{\nu}$" is the root-mean-square error between the NLDC and its approximation; "$NRMSE^{\nu}$" is the normalized root-mean-square error between the NLDC and its approximation.

2.5.7 Commitment Results and Errors

Despite the small error incurred by the four-week approximation in representing the net load, its commitment performance must also be tested. This test is based on comparing the commitment results obtained from running IMRES with the one-week, two-week, four-week and the eight-week approximation, with those from a UC model that uses the full one-year data series.

The test is performed with data corresponding to the ERCOT system, allowing power curtailment of both wind and solar PV, and using standard thermal units whose cost parameters and technical characteristics can also be found in the Appendix. The solver chosen for this experiment is CPLEX 12.4 under GAMS on a 64-bit dual-socket hexcore (i.e. 12 physical cores, 12 virtual cores total) Intel Nehalem (x5650) machine. The maximum duality gap tolerance was set to 0.01, producing a solution in approximately 4 hrs. Tighter duality gap tolerances did not affect the main solution parameters but had a significant impact on the computation time. The computing time with commercial machines (2.4 GHz Intel Core 2 Duo machine with 4GB of memory) was on the order of 1.5 times greater than those experienced with the original 12-core machine. Panel A) in Tables 2.7-2.10 summarizes the capacity and commitment results obtained with each approximation (each IMRES

simulation uses a different NLDC approximation in its unit commitment component).

The second component of the test consists of running a full one-year UC model for each of the capacity mixes produced by IMRES. The full year UC model uses the same parameters and applies the same constraints as those in IMRES, but uses hourly demand and renewable data for the 8 760 hours of the year. In order to test all the approximations, the full year UC is performed for the different capacity mixes determined by each IMRES simulation. The comparison is then established between the commitment results derived from the full year and those derived from the approximation, both using the same capacity mix. Panel B) in Tables 2.7-2.10 presents the results from the full year run and the differences with the results derived using the approximations tested.

Table 2.7: One-Week Net-Load-Based Approximation Results and Errors

A) One-Year Unit Commitment Results Based on the One-Week Net-Load-Based Approximation.

Technology	Units Installed	Capacity Installed [GW]	EC [p.u.]	CF [p.u.]	Startups/yr
Nuclear	8	8	0.245	1	0
Solar	10	10	0.043	0.141	-
Wind	30	30	0.190	0.206	-
Coal	13	6.5	0.188	0.943	0
CCGT	72	28.8	0.331	0.375	102
OCGT	18	5.4	0.003	0.017	145

B) One-Year Unit Commitment Results Based on a Full One-Year Data Series and Commitment Error.

Technology	EC [p.u.]	ECE [p.u.]	CF [p.u.]	CFE [p.u.]	Startups/yr	SUE [startups/yr]
Nuclear	0.228	+0.017	1	+0.000	0	+0
Solar	0.056	-0.013	0.198	-0.054	-	-
Wind	0.199	-0.009	0.233	-0.027	-	-
Coal	0.176	+0.012	0.951	-0.008	0	+0
CCGT	0.325	+0.006	0.395	-0.020	64	+38
OCGT	0.007	-0.004	0.048	-0.031	148	-3

Result summary from: A) a capacity expansion model with embedded unit commitment constraints based on a one-week sample; B) a one-year full unit commitment model of the units installed, and commitment errors between the full run and the approximation. NSE = 7.75E-3 p.u.

Table 2.8: Two-Week Net-Load-Based Approximation Results and Errors

A) One-Year Unit Commitment Results Based on the Two-Week Net-Load-Based Approximation.

Technology	Units Installed	Capacity Installed [GW]	EC [p.u.]	CF [p.u.]	Startups/yr
Nuclear	9	9	0.243	1	0
Solar	10	10	0.058	0.215	-
Wind	30	30	0.226	0.279	-
Coal	12	6	0.154	0.950	0
CCGT	87	34.8	0.317	0.338	74
OCGT	18	5.4	0.002	0.014	97

B) One-Year Unit Commitment Results Based on a Full One-Year Data Series and Commitment Error.

Technology	EC [p.u.]	ECE [p.u.]	CF [p.u.]	CFE [p.u.]	Startups/yr	SUE [startups/yr]
Nuclear	0.257	-0.014	1	+0.000	0	+0
Solar	0.057	+0.001	0.198	+0.017	-	-
Wind	0.196	+0.030	0.229	+0.050	-	-
Coal	0.162	-0.008	0.950	+0.000	0	+0
CCGT	0.325	-0.008	0.327	+0.011	57	+17
OCGT	0.003	-0.001	0.019	-0.005	97	+0

Result summary from: A) a capacity expansion model with embedded unit commitment constraints based on a two-week sample; B) a one-year full unit commitment model of the units installed, and commitment errors between the full run and the approximation. NSE = 7.75E-3 p.u.

Table 2.9: Four-Week Net-Load-Based Approximation Results and Errors

A) One-Year Unit Commitment Results Based on the Four-Week Net-Load-Based Approximation.

Technology	Units Installed	Capacity Installed [GW]	EC [p.u.]	CF [p.u.]	Startups/yr
Nuclear	8	8	0.226	1	0
Solar	10	10	0.063	0.223	-
Wind	30	30	0.198	0.233	-
Coal	15	7.5	0.199	0.939	0
CCGT	87	34.8	0.310	0.315	54
OCGT	30	9	0.003	0.012	61

B) One-Year Unit Commitment Results Based on a Full One-Year Data Series and Commitment Error.

Technology	EC [p.u.]	ECE [p.u.]	CF [p.u.]	CFE [p.u.]	Startups/yr	SUE [startups/yr]
Nuclear	0.228	-0.002	1	+0.000	0	+0
Solar	0.057	+0.007	0.199	+0.024	-	-
Wind	0.199	-0.001	0.233	+0.000	-	-
Coal	0.199	+0.000	0.933	+0.006	0	+0
CCGT	0.313	-0.003	0.316	-0.001	59	-5
OCGT	0.003	+0.000	0.012	+0.000	65	-4

Result summary from: A) a capacity expansion model with embedded unit commitment constraints based on a four-week sample; B) a one-year full unit commitment model of the units installed, and commitment errors between the full run and the approximation. NSE = 7.75E-3 p.u.

Table 2.10: Eight-Week Net-Load-Based Approximation Results and Errors

A) One-Year Unit Commitment Results Based on the Eight-Week Net-Load-Based Approximation.

Technology	Units Installed	Capacity Installed [GW]	EC [p.u.]	CF [p.u.]	Startups/yr
Nuclear	10	10	0.284	1	0
Solar	10	10	0.063	0.220	-
Wind	30	30	0.186	0.219	-
Coal	11	5.5	0.148	0.950	0
CCGT	89	35.6	0.316	0.312	66
OCGT	26	7.8	0.003	0.012	71

B) One-Year Unit Commitment Results Based on a Full One-Year Data Series and Commitment Error.

Technology	EC [p.u.]	ECE [p.u.]	CF [p.u.]	CFE [p.u.]	Startups/yr	SUE [startups/yr]
Nuclear	0.285	-0.001	1	+0.000	0	+0
Solar	0.056	+0.007	0.197	+0.023	-	-
Wind	0.192	-0.006	0.225	-0.006	-	-
Coal	0.149	-0.001	0.950	+0.000	0	+0
CCGT	0.315	+0.001	0.310	+0.002	57	+9
OCGT	0.002	+0.001	0.011	+0.001	67	+4

Result summary from: A) a capacity expansion model with embedded unit commitment constraints based on an eight-week sample; B) a one-year full unit commitment model of the units installed, and commitment errors between the full run and the approximation. NSE = 7.75E-3 p.u.

A comparison of the errors produced by the different sample sizes, shows that the errors corresponding to the one-week and two-week sample runs are greater than those found in the four- week and eight-week sample runs. Both energy contribution errors and capacity factor errors are at least twice as large in the first two cases as in the last two. Also, start-up errors in the first two cases are close to or well over 20 cycles per year, indicating that they represent poorly the cycling behavior of the thermal units, a main concern of flexibility studies. On the other hand, the average number of start-ups per year in the four-week and eight-week sample are very similar to the results of the full-year UC simulation.

Additionally, the quantity of NSE in the four-week and eight-week cases are one order of magnitude below those produced by the one-week and the two-week sample. This difference indicates that the larger sample sizes produce a more adequate capacity mix than the first two because the variation in net load is better represented. Errors are not appreciably reduced by using a sample size of eight weeks, relative to the four-week approximation.

Based on these results, an approximation of the net load from four-weeks of data using this approach appears to represent well the full one-year data series. The optimal four-week approximation of net load will be used for the remainder of this paper.

2.6 Representing the Peak Net Load

One of the critical parameters for determining adequacy in capacity expansion and flexibility assessment studies is the value-of-lost-load (VOLL) (Stoft, 2002). An adequate system typically has less than one day with non-served energy (NSE) every 10 years (or less than 2.4 hours every year), although it varies from one system to another. In this thesis the VOLL is assumed to be fixed to \$500/MWh (see section 2.4.3), and the amount of NSE is derived endogenously from the cost minimization model. Nevertheless, the amount of NSE does not depend exclusively on the VOLL. It is also determined by the system's peak load or, in situations with a high penetration of renewables, the peak net load. Hence, it is important to examine how including the peak net load in the sample affects the final solution of the problem.

Comparing the four-week approximation ν^* with the NLDC, the approximation underestimates the peak hour by 4.5%. If embedded within a capacity expansion model, this underestimate would lead to insufficient investment in generation capacity.

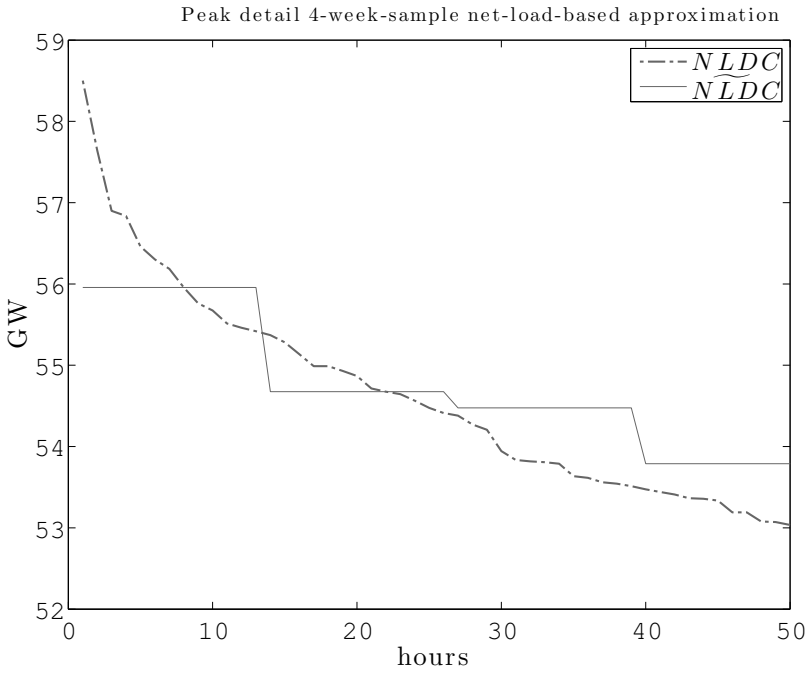


Figure 2-4: Detail of the system's NLDC peak and its respective four-week-based approximation. The peak value in both series is 68.5 GW and 56 GW respectively. Note that since the number of weeks selected is four, the step size of the $\widetilde{NLDC}$ is $\Theta = 52/4 = 13$.

One possible solution to this issue is to require having the week containing the peak net load as one of the weeks in the sample. We denote this approximation by κ^*. The κ^* approximation is constructed with the week containing the peak net load plus an optimal selection of three weeks, conditional on having the peak net load week in the approximation.

Alternatively, we can revisit the initial assumption that the year is composed of 52 weeks (in reality it is composed of 52 weeks and one day), and add the peak net load day to one four-week approximation, to make a total of 8,760 hours in one year. This approximation will be denoted by τ^*. Figure 2-5 shows the detail of the peak net load as achieved by these two alternative approximations that contain the peak net load day.

The accuracy of each approximation can be assessed by the NRMSE metric and by the ca-

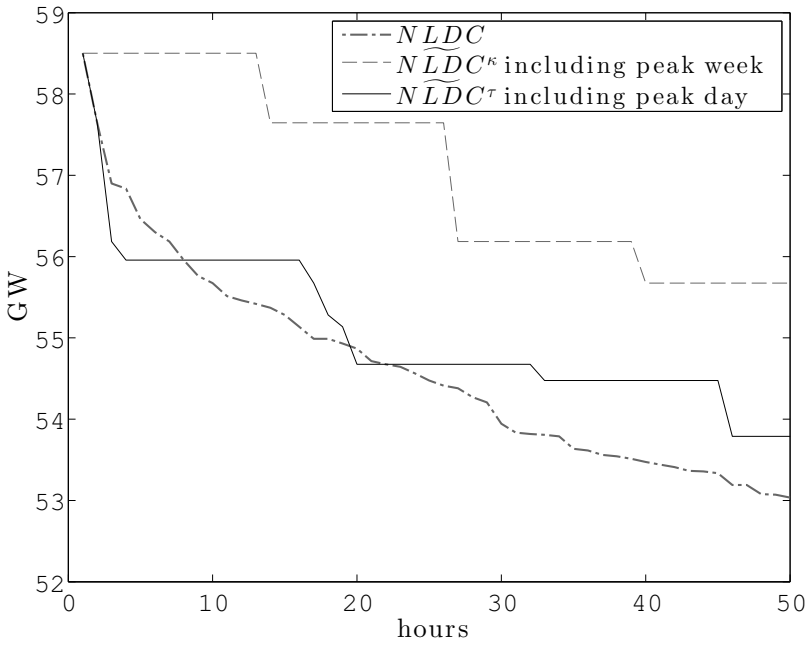


Figure 2-5: Detail of the system's NLDC peak and its κ^* and its τ^* four-week-based approximations. The peak value in both series is 68.5 GW and 56 GW respectively. Note that since the number of weeks selected is four, the step size of the $\widetilde{NLDC}$ is $\Theta = 52/4 = 13$.

pacity results obtained from running IMRES with each of the three different approximations (Table 2.11).

The results show that approximations τ^* and κ^*, with greater peak capacities and lower NSE, represent the peak net load more precisely than the ν^* approximation. However, the NRMSE is greater for the approximation κ^* because it is forcing a suboptimal week to be included in the sample. This error is reduced with approximation τ^* because the optimal selection is done conditional on having introduced the peak net load day. As a result of better resolution in the peak of the NLDC, approximations that consider the peak net load yield a higher system capacity and lower NSE. Although the impact on the overall approximation of the NLDC is small, including the peak net load day in the approximation increases the accuracy in the peak hours.

Table 2.11: Standard Error Relative to the Number of Weeks Selected

Number of Weeks	No Peak (ν^*)	Peak Week (κ^*)	Peak Day (τ^*)
Weeks Selected	{17,31,37,46}	{29,41,46,48}	{17,21,31,45} + peak net load day (day #197)
NRMSE [%]	0.546	0.821	0.537
Total Capacity Installed [GW]	99.3	102.3	99.4
NSE [p.u.]	3.79E-4	9.74E-5	1.24E-4
System Cost [%]	-	-0.54	-1.86

Summary of the error sum of squares incurred by three approximations: 1) neglecting the peak net load; 2) including the week containing the peak net load hour (52 weeks); and 3) including the day containing the peak net load hour (52 weeks + 1 day). The peak net load hour is in day #197, in week #29.

2.7 Comparison to Other Approaches

This section shows the advantages of the net-load-based selection approach by comparing it with other sample week selection methods found in the literature.

2.7.1 Results from Using a Season-Based Selection

The most common approach to selecting a sample of weeks to approximate the demand in one year is to identify seasonal demand patterns, and sampling one week from each of the seasons (Kirschen et al., 2011). For example, the maximum, average and minimum demand from each week can be used to identify the four seasons. One week from each season is chosen randomly to make up a sample of four weeks, and the hours in the sample are scaled up to construct a one-year season-based approximation ρ^*. For this example, the sample $\rho^* = \{14, 27, 40, 49\}$ has a NRMSE between the NLDC and its approximation of 2.78, as compared with 0.5 from our four-week approximation. Table 2.12 shows the capacity and commitment results using a season-based approximation, and commitment results with the three error metrics.

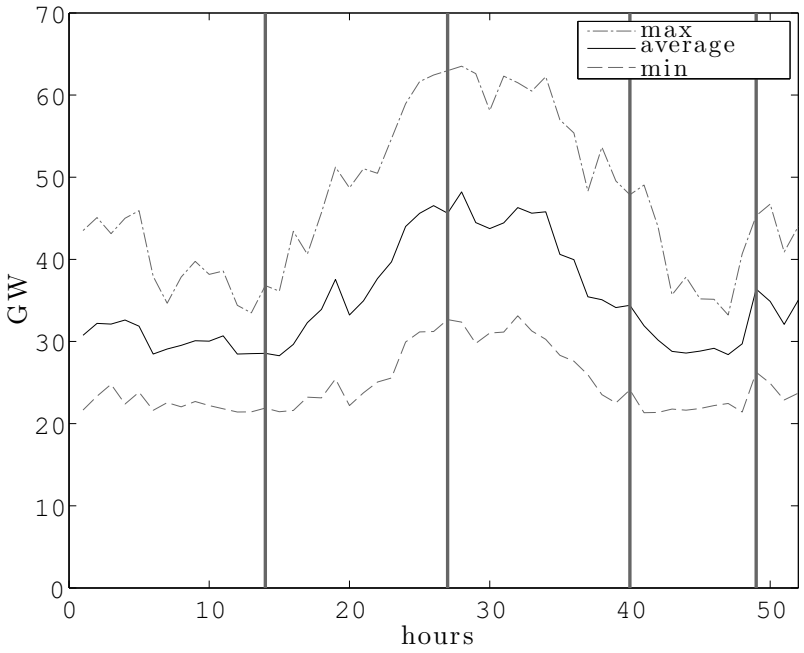


Figure 2-6: Graphical representation of the maximum, average and minimum weekly demand for the ERCOT system throughout 2008.

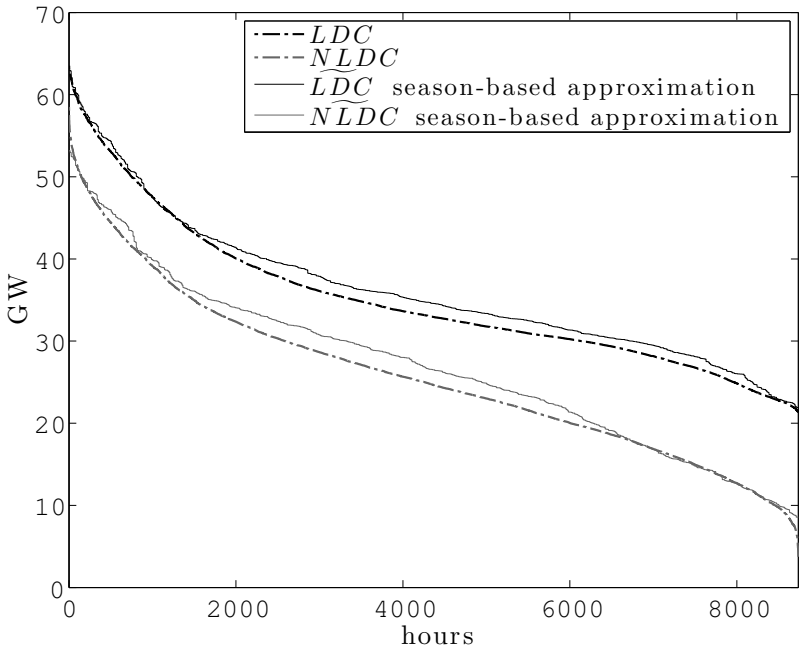


Figure 2-7: Graphical representation of the system's LDC, NLDC, and their respective four-week approximations selected through seasonality criteria.

Table 2.12: Four-Week Net-Season-Based Approximation Results and Errors

A) One-Year Unit Commitment Results Based on the Four-Week Season-Based Approximation.

Technology	Units Installed	Capacity Installed [GW]	EC [p.u.]	CF [p.u.]	Startups/yr
Nuclear	11	11	0.294	1	0
Solar	10	10	0.046	0.171	-
Wind	30	30	0.209	0.260	-
Coal	14	7	0.173	0.924	0
CCGT	88	35.2	0.276	0.293	76
OCGT	29	8.7	0.003	0.014	80

B) One-Year Unit Commitment Results Based on a Full One-Year Data Series and Commitment Error.

Technology	EC [p.u.]	ECE [p.u.]	CF [p.u.]	CFE [p.u.]	Startups/yr	SUE [startups/yr]
Nuclear	0.314	-0.020	1	+0.000	0	+0
Solar	0.056	-0.010	0.196	-0.025	-	-
Wind	0.189	+0.020	0.221	+0.039	-	-
Coal	0.180	-0.007	0.900	+0.024	0	+0
CCGT	0.260	+0.017	0.259	+0.034	68	+8
OCGT	0.002	+0.001	0.006	+0.008	53	+27

Result summary from: A) a capacity expansion model with embedded unit commitment constraints based on an four-week sample; B) a one-year full unit commitment model of the units installed, and commitment errors between the full run and the approximation. NSE = 3.65E-3 p.u.

2.7.2 Results from Ignoring the Correlation Between Renewables and Load

The season-based selection approach used above has two sources of error: first, it randomly chooses a week from each season rather than building the selection upon any error minimization criterion; and second, it focuses on load rather than net load, ignoring the correlation between demand and renewable output. In order to differentiate between these two sources of error, we show the errors derived from using an optimal selection procedure, but aimed at approximating the LDC, which ignores the correlation between load and renewables. Analogously to the optimization problem proposed in (2.51), the objective of the selection problem now is to find the week combination that yields the least square error between the LDC and its approximation, $\widetilde{LDC}$ (2.57).

$$\lambda^* = \arg\min_{\lambda} \|LDC - \widetilde{LDC^{\lambda}}\|^2 \tag{2.57}$$

–where $\lambda \in \mathbb{Z}^n$ is the set of indices of the n weeks selected; $\lambda^* \in \mathbb{Z}^n$ is the set of indices of the optimal week combination; and $LDC, \widetilde{LDC}^{\nu} \in \mathbb{R}^{8,736}$.

This problem is solved through enumeration, by constructing all possible approximations for a certain number of weeks, and selecting the combination of weeks that best represents the LDC. Accordingly, the $\widetilde{NLDC}^{\lambda}$ is built with the net load associated with the weeks selected through this process.

For all cases explored (one, two, four, eight, twelve and sixteen week samples), the resulting error metrics are summarized in Table 2.13

The errors from all the approximations corresponding to different number of weeks are significantly greater when the selection is based on the LDC (λ^*) than when the selection is based on approximating the $NLDC$ (ν^*). By fitting the LDC, we do not necessarily select the weeks that characterize best the $NLDC$ or, in other words, a selection based on the LDC is suboptimal relative to how well it characterizes the $NLDC$.

More importantly, the errors in the approximate capacity, commitment, and startup

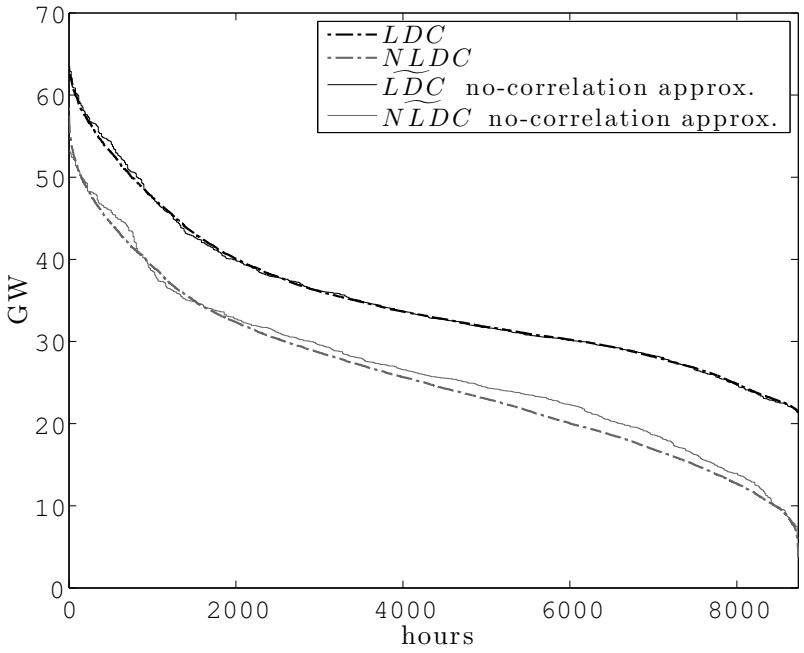


Figure 2-8: Graphical representation of the system's LDC, NLDC and their respective four-week approximations, selected to fit the LDC.

results obtained with λ^* are larger when the sample is based on the LDC (Table 2.14). As the level of penetration of renewables increases, so does the importance of accounting for the correlation between load and renewable output.

Table 2.13: Net Energy Error Relative to the Number of Weeks Selected

Number of Weeks	Approximation λ*	$RMSE^{\lambda}$	$NRMSE^{\lambda}$ [%]	time [sec]
1	{38}	3.953	7.22	0.06
2	{10,23}	2.311	4.22	1.61
4	{11,16,27,40}	1.271	2.32	694
8	{5,6,16,17,24,25,39,40}	1.242	2.27	992
12	{4,5,6,24,25,26,39,40,41,42,43,44}	0.627	1.14	1,419
16	{3,4,5,6,15,16,17,18,24,25,26,27,39,40,41,42}	0.583	1.07	1,542

Summary of the error incurred by approximations constructed with an increasing number of weeks. The weeks are selected based on minimizing the error between the approximation and the LDC. The column "Approximation" shows the weeks selected; "$RMSE^{\lambda}$" is the root-mean-square error between the NLDC and its approximation; "$NRMSE^{\lambda}$" is the normalized root-mean-square error between the NLDC and its approximation.

Table 2.14: Four-Week Load-Based Approximation Results and Errors

A) One-Year Unit Commitment Results Based on the Four-Week Load-Based Approximation.

Technology	Units Installed	Capacity Installed [GW]	EC [p.u.]	CF [p.u.]	Startups/yr
Nuclear	10	10	0.285	1	0
Solar	10	10	0.062	0.218	-
Wind	30	30	0.162	0.190	-
Coal	14	7	0.185	0.928	0
CCGT	86	34.4	0.303	0.309	55
OCGT	21	6.3	0.002	0.012	69

B) One-Year Unit Commitment Results Based on a Full One-Year Data Series and Commitment Error.

Technology	EC [p.u.]	ECE [p.u.]	CF [p.u.]	CFE [p.u.]	Startups/yr	SUE [startups/yr]
Nuclear	0.285	+0.000	1	+0.000	0	+0
Solar	0.056	+0.006	0.197	+0.021	-	-
Wind	0.193	-0.031	0.225	-0.035	-	-
Coal	0.183	+0.002	0.915	+0.013	0	+0
CCGT	0.281	-0.002	0.286	+0.023	65	-10
OCGT	0.002	+0.000	0.012	+0.000	81	-12

Result summary from: A) a capacity expansion model with embedded unit commitment constraints based on an four-week sample; B) a one-year full unit commitment model of the units installed, and commitment errors between the full run and the approximation. NSE = 3.98E-3 p.u.

2.8 Validity Assessment

In order to show the applicability of the method to other power systems, this section tests the four-week sampling method described in previous sections with data (load, wind CF and solar CF) corresponding to Germany during year 2008. The resulting four-week sample is then used to build an approximation of the net load, which serves as input to IMRES. In the same way as with the tests performed above, a full-year unit commitment of the capacity mix obtained is used to assess the quality of the net load approximation in terms of representing the chronological variability of the net load. Results are shown in Table 2.15 using the same format as in previous simulations.

The magnitude of the commitment errors are similar to the errors obtained with the ERCOT 2009 system. This result gives additional confidence in the applicability of the approximation method to power systems with a high penetration of intermittent renewables. The results presented here taken together suggest that our approximation method might perform well for a range of systems.

Table 2.15: Four-Week Net-Load-Based Approximation Results and Errors

A) One-Year Unit Commitment Results Based on the Four-Week Load-Based Approximation.

Technology	Units Installed	Capacity Installed [GW]	EC [p.u.]	CF [p.u.]	Startups/yr
Nuclear	30	30	0.530	1	0
Solar	10	10	0.019	0.106	-
Wind	30	30	0.105	0.198	-
Coal	16	8	0.130	0.924	0
CCGT	78	31.2	0.215	0.390	68
OCGT	17	5.1	0.001	0.015	97

B) One-Year Unit Commitment Results Based on a Full One-Year Data Series and Commitment Error.

Technology	EC [p.u.]	ECE [p.u.]	CF [p.u.]	CFE [p.u.]	Startups/yr	SUE [startups/yr]
Nuclear	0.532	-0.002	1	+0.000	0	+0
Solar	0.017	+0.002	0.096	+0.010	-	-
Wind	0.102	+0.003	0.192	+0.006	-	-
Coal	0.131	-0.001	0.926	-0.002	0	+0
CCGT	0.216	-0.001	0.391	-0.001	76	-8
OCGT	0.001	+0.000	0.013	+0.002	99	-2

Result summary from: A) a capacity expansion model with embedded unit commitment constraints based on an four-week sample; B) a one-year full unit commitment model of the units installed, and commitment errors between the full run and the approximation. NSE = 3.92E-3 p.u.

2.9 Conclusions

This chapter introduced IMRES, a capacity expansion model with unit commitment constraints embedded that incorporates the extra costs of cycling and operational constraints into investment decisions. In addition to the model, the chapter also presents a methodology to sample weeks based on historical demand and renewable generation data, that can be used to reduce the dimensionality of capacity expansion models with unit commitment constraints such as the one use in this thesis. IMRES provides an effective approach for reducing the computation time for capacity expansion with embedded unit commitment constraints, while still capturing the relevant characteristics of a system with renewables. Given the increased concern about flexibility in future generation planning, this approach provides a means to evaluate the flexibility of alternative future generation mixes.

The analyses performed for different number of weeks illustrate how the error of the approximation depends on the number of weeks included, and that the error incurred with a four-week approximation is small with respect to both net energy and unit commitment error metrics. In addition, comparisons with other methods in the literature indicate that in order to best reflect the possible correlation between demand and renewables, the selection should be based on the net load instead of attending to other aspects such as demands seasonal patterns or the LDC. Lastly, the similarity of the results obtained from applying the selection method to a different power system helps to confirm the consistency of the results.

Nevertheless, the week selection method proposed only guarantees that the four weeks selected represent well the energy below the net load duration curve, without explicitly accounting for the chronological variability of the net load. Therefore, it is recommended that any specific application of this methods performs the full-year unit commitment test with the capacity resulting from IMRES (focusing on the three commitment error metrics), to guarantee that the $\widetilde{NLDC}$ represents as well the chronological variability of the $NLDC$. In the future, other improved selection criteria based on commitment metrics could be explored to avoid this caveat.

2.10 Applications

The formulation of IMRES offers a great versatility in terms of which constraints are included in the model, which are left out, and which decisions variables are free and which ones are fixed and parameterized. Therefore, IMRES could serve as modeling framework for a large set of experiments. As an example of the potential capabilities of IMRES, the experiments performed in IEA's report *The Power of Transformation. Wind, Sun and the Economics of Flexible Power Systems* (IEA, 2014) include a comprehensive economic analysis of the impact of intermittent renewables on power systems and a cost-benefit analysis of three different flexibility options (demand side management, storage and thermal power plants). As a summary, the following list presents a sample of possible tasks that could be performed with IMRES:

- Determining the optimal electricity generation capacity mix at a given future time.
- Analyzing the effect of constraints (such as total carbon emissions) on the total system cost through selectively applying and removing these constraints.
- Calculating the value of *flexibility* options (storage, DSM, flexible power plants, etc) by comparing the total system cost with and without each option, to establish R&D priorities.
- Calculating the profitability of individual units through obtaining the prices and the quantities produced for the different services provided.
- Analyzing the effect of different renewable deployment levels for a fixed capacity mix.

Chapter 3

The Impact of Intermittent Renewables on Electricity Markets

This chapter is an extended adaptation of the following paper:

- "The Impact of Bidding Rules on Electricity Markets with Intermittent Renewables", submitted to *IEEE Transactions on Power Systems*, IEEE (de Sisternes et al., 2014)

3.1 Introduction

The operation of power systems with large amounts of intermittent renewables is expected to change substantially from how the system operates solely with conventional technologies: first, large shares of zero-variable-cost renewables reduce the peak and change the shape of the net load (the load that has to be supplied by conventional technologies); second, renewables' variability make conventional technologies to cycle more frequently in order to maintain the instantaneous balance between supply and demand; and third, renewables' uncertainty increase the need for reserve capacity and its utilization. This chapter explores the impacts of renewables on electricity markets by focusing on the prices[1] and profits obtained by plants

[1]since the trade of electricity is typically referenced to the price of the day-ahead market, this chapter focuses exclusively on this market, leaving the study of the implications on secondary and reserve markets for further research.

under different market conditions.

Renewable integration studies (Holttinen et al., 2008)(Perez-Arriaga and Batlle, 2012)(Troy et al., 2010)(Delarue et al., 2007) have shown that cycling of thermal power plants significantly intensifies in the presence of intermittent renewables, increasing fuel and operation & maintenance costs from startup of the thermal units. In this context, it is particularly relevant to study the mechanisms through which thermal power plants participating in liberalized markets with renewables recover the costs associated with starting up and how these mechanisms affect the wholesale electricity price and the total cost recovery [2]. Ultimately, the profitability of the generating units in the system can be examined.

Many wholesale markets have adopted complex bidding mechanisms to account for non-convex cost components (those associated with discrete decisions, such as start up or the minimum output requirement) that are not reflected in the marginal cost of energy (O'Neill et al., 2005). Such complementary bidding procedures for instance allow companies to declare their startup costs or to specify a minimum revenue threshold for the day that ensures the cost-recovery of the units. Examples of these different procedures include an energy-based market price where startup costs are paid separately for bids that are accepted (similarly to the PJM market (Sioshansi, 2008)); a market price that includes an uplift added to the energy component of price to guarantee the recovery of operating costs (comparable to the Irish market (AIP, 2007)); and internalizing startup costs in the bid into a single energy term (in a similar way as it is done in general in other European markets (Rothwell and Gomez, 2003)). This paper uses a capacity expansion model with embedded unit commitment constraints to explore the different implications of each bidding rule for the resulting wholesale price, the cost recovery by individual units, and the cost borne by consumers.

Concurrently with the impact of pricing rules for recouping startup costs on market prices, the regulatory uncertainty around renewable deployment has gained significant attention as a critical factor affecting the profitability of thermal power plants. In most cases,

[2]Remember that in this thesis by "profit" it is understood any excess income to a plant from market prices that is beyond strict cost recovery: operation costs (here only startup costs and fuel costs of electricity production have been considered), plus the capital costs (amortization and a reasonable rate of return on investment).

renewable deployment still depends on the existence of subsidies (e.g., feed-in-tariffs or tax credits) or regulatory mandates (e.g., renewable portfolio standards). Unforeseen regulatory changes that affect the deployment pace of renewables can alter the economic equilibrium of the generation fleet. Such regulatory shocks, combined with the short construction time of renewable projects, can quickly displace some thermal generators in the merit order, potentially resulting in partly or totally stranded assets among the thermal generation fleet. Our analysis also studies the relative importance of these two factors – bidding rules and regulatory uncertainty and distinguishes between their effects.

Much of the work found in the literature on this topic has been devoted to investigating the short-term impact of renewables on electricity markets, focusing on determining how greater shares of renewables will change wholesale prices and the remuneration of existing plants (Poyry and EWEA, 2010)(Brown, 2012), or on how new investments could be incentivized (Newell, 2012)(Baritaud, 2012). Other efforts have been focused on calculating the long-term market equilibrium capacity mix, or quantifying the benefits of elements such as storage or demand response in systems with a high share of renewables (de Jonghe et al., 2011)(Palmintier and Webster, 2011). Yet, little attention has been paid to studying the interaction between each one of the different bidding mechanisms found in wholesale markets, the renewable penetration level, and the economic losses occurring in a generating fleet that is ill-adapted following an unanticipated addition of renewable capacity to the system. This chapter fills this gap by unpacking the interaction between these three elements, explaining their relative influence on the wholesale clearing price and the remuneration of four specific thermal generation technologies: nuclear, coal, combined cycle gas turbines (CCGTs), and open cycle gas turbines (OCGTs).

The chapter is structured as follows. First, we present the experimental design used in the study, highlighting the differences between an adapted and a non-adapted system. Second, we describe the method used in the experiments to determine the minimum cost generation fleet, the dispatch, the hourly prices, and the remuneration of the units. We also describe the implementation of the bidding rules studied and the calculation of the

plants' profits for each rule. Third, we propose a heuristic method that selectively adds or removes plants based on their expected remuneration to derive an approximate equilibrium solution for the generation fleet, used when the bidding rules overpay or underpay certain technologies. Finally, we present the results of the simulations, show the markedly different implications to generators and to consumers of the bidding rules studied, and present the conclusions.

3.2 Experimental Design

The proposed experimental design allows for a comparative analysis of the effect of bidding rules under ten different renewable penetration scenarios. In each scenario (S), wind capacity and solar PV capacity have both been chosen to be equal to $10 \cdot S$ gigawatts, where $S \in \{1, \ldots, 10\}$. Although the capacity installed of each technology is the same, their energy contributions will differ, as wind and solar PV have different capacity factors and production patterns.

From optimal marginal pricing theory (Perez-Arriaga and Meseguer, 1997), if technologies are able to adapt and the ideal conditions are satisfied, in equilibrium, and in the absence of common or internal constraints on installed capacity, the profit of all power plants will be exactly zero. Alternatively, if there is a renewable deployment shock in the system, the ensuing excess capacity situation will produce a transition period during which some units in the generation fleet will likely incur losses. In equilibrium, generation is perfectly adapted to renewable deployment and remuneration gaps are exclusively due to the effect of nonconvexities, which allows for the assessment of the performance of bidding rules in recovering nonconvex costs. Without equilibrium, the effect of the nonconvexities interacts with the effect of having non-adapted system. This situation can be useful to estimate the relative effect of the bidding rules under the influence of an unexpected deployment of intermittent renewables. We therefore consider two cases:

- **Adaptation**: represents market equilibrium conditions, as simulated by the solution of

a capacity expansion model that chooses both investment and operations, accounting for unit commitment constraints and minimizes total system costs. The result of this exercise is a generation fleet $\boldsymbol{y}^*(S)$ that is perfectly adapted to a given renewable capacity S.

- **Non-adaptation**: assumes the generation mix that minimizes costs for zero renewable capacity, which then operates in the presence of a given capacity S of renewable generation, $\boldsymbol{y}^0(S)$.

Throughout the analysis we assume that there is no market power and that a cost minimization approach can be used to obtain a market equilibrium solution. Once we have the capacity and operational variables from both cases, we use a market simulator that calculates the wholesale prices under a particular bidding rule (see Section V) and the cash flows for all the units in the system. The final output of the market simulator, for both the adapted and the non-adapted case, is the average profit for each technology and the average cost to consumers. The resulting experimental design is depicted in Fig. 3-1.

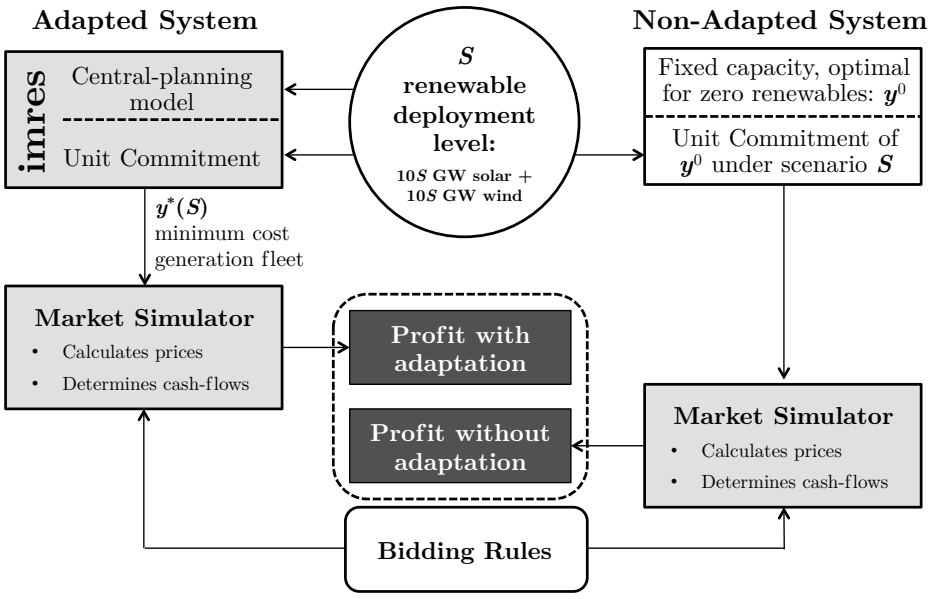


Figure 3-1: Experimental design

Different bidding rules will result in a different balance between producer and consumer

surplus. Therefore, they have to be assessed on the basis of how they affect these two types of agents, measured through two separate metrics:

1. *Average profit* of technology k ($\bar{\pi}_k$), obtained by a thermal technology and reflects the extent to which the remuneration obtained by the units exceeds their fixed and variable costs (3.1)

$$\bar{\pi}_k = \frac{\sum_{i \in \mathcal{T}^k} \pi_i}{\sum_{i \in \mathcal{T}^k} y_i} \qquad \forall k \in \mathcal{K} \mid \sum_{i \in \mathcal{T}^k} y_i > 0 \tag{3.1}$$

2. *Average cost to the consumer* ($\overline{CC}$)[3], accounts for the average price paid by consumers, resulting from dividing all the costs paid by the consumer (excluding those derived from incentive plans to renewable technologies) by the total energy consumed (3.2).

$$\overline{CC} = \frac{\sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} \rho_j x_{ij} + \sum_{i \in \mathcal{I}} \sum_{d \in \mathcal{D}} s_{id}}{\sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} x_{ij}} \tag{3.2}$$

3.3 Methodology

The operational regime of thermal plants becomes more demanding with greater renewable deployment as a consequence of a more variable net load. The technical constraints that were typically ignored in generation planning models (Turvey and Anderson, 1977) (minimum required output of the units, output ramping limits, minimum up and down times, and startup costs) become binding and must be incorporated into the model to account for their effect on the system cost.

The analyses in this chapter use IMRES, a centralized capacity expansion model with detailed unit commitment constraints embedded (de Sisternes and Webster, 2013), to account for the impact of nonconvex costs on investment decisions. In each case, the model chooses both the investments in new capacity needed in the system and the hourly commitment and

[3]Note that in this expression the side payment s_{id} takes on non-zero values only when the bidding rule involving side payments is analyzed.

generation schedule of all units to meet demand at minimum cost. The objective function in IMRES minimizes the total cost of the system, which is the sum of investment and operating costs, including the cost of non-served energy (3.3). As explained in the previous chapter, a set of four representative weeks in the unit commitment component of the model that approximate the inter-hour variability of the net load over the year to reduce the dimensionality of the problem (de Sisternes, 2013).

$$\min_{\boldsymbol{x},\boldsymbol{y},\boldsymbol{z},\boldsymbol{u}} \quad \sum_{i\in\mathcal{I}} C_i^{FIX} y_i + \Theta \left(\sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}} (C_i^{VAR} x_{ij} + C_i^{STUP} z_{ij}) + \sum_{j\in\mathcal{J}} VOLL\ n_j \right) \tag{3.3}$$

The analyses in this thesis focus on the day-ahead market, as electricity trade at different time-scales is referenced to this market. Therefore, all constraints relative to reserves have been ignored in the model, to avoid introducing constraints that are not remunerated. The wholesale energy price is determined based on the dual variable of the demand balance equation and the bid-forming mechanisms implicit in each of the rules described below.

3.3.1 Operational Constraints

The model includes the most important operation constraints in unit commitment formulations: demand balance equation (3.25); commitment and maximum and minimum output constraints minimum output requirement (3.5-3.8); maximum ramp constraints (3.9-3.10); constraints representing the minimum up and down time of the plants (3.11-3.12); constraints setting the renewable scenarios (3.13-3.14); and constraints allowing for the curtailment of renewable energy output (3.15-3.16). In addition, the model includes a coupling constraint that imposing the condition that only units that have been built can generate (3.17).

$$\sum_{i\in\mathcal{I}} x_{ij} + n_j = D_j \qquad \forall j \in \mathcal{J} \tag{3.4}$$

$$u_{ij} - u_{ij-1} = z_{ij} - v_{ij} \qquad \forall i \in \mathcal{T},\ \forall j \in \mathcal{J}\backslash\{1\} \tag{3.5}$$

$$z_{ij} + v_{ij} \leqslant 1 \qquad \forall i \in \mathcal{T},\ \forall j \in \mathcal{J} \tag{3.6}$$

$$w_{ij} = x_{ij} - u_{ij}\ P_i^{MIN} \qquad \forall i \in \mathcal{T},\ \forall j \in \mathcal{J} \tag{3.7}$$

$$w_{ij} \leqslant u_{ij}(P_i^{MAX} - P_i^{MIN}) \qquad \forall i \in \mathcal{T},\ \forall j \in \mathcal{J} \tag{3.8}$$

$$w_{ij} - w_{ij-1} \leqslant R_i^U \qquad \forall i \in \mathcal{T},\ \forall j \in \mathcal{J}\backslash\{1\} \tag{3.9}$$

$$w_{ij-1} - w_{ij} \leqslant R_i^D \qquad \forall i \in \mathcal{T},\ \forall j \in \mathcal{J}\backslash\{1\} \tag{3.10}$$

$$u_{ij} \geqslant \sum_{j'>j-M_i^U}^{j} z_{tj'} \qquad \forall i \in \mathcal{T},\ \forall j \in \mathcal{J} \tag{3.11}$$

$$1 - u_{ij} \geqslant \sum_{j'>j-M_i^D}^{j} v_{tj'} \qquad \forall i \in \mathcal{T},\ \forall j \in \mathcal{J} \tag{3.12}$$

$$\sum_{i\in\mathcal{W}} y_i = 10\ S \tag{3.13}$$

$$\sum_{i\in\mathcal{S}} y_i = 10\ S \tag{3.14}$$

$$x_{ij} \leqslant P_i^{MAX}\ CF_j^{WIND} \qquad \forall i \in \mathcal{W},\ \forall j \in \mathcal{J} \tag{3.15}$$

$$x_{ij} \leqslant P_i^{MAX}\ CF_j^{SOLAR} \qquad \forall i \in \mathcal{S},\ \forall j \in \mathcal{J} \tag{3.16}$$

$$x_{ij} \leqslant P_i^{MAX}\ y_i \qquad \forall i \in \mathcal{I},\ \forall j \in \mathcal{J} \tag{3.17}$$

Non-negativity constraints and integrality constraints (3.18-3.24):

$$x_{ij} \geqslant 0 \qquad \forall i \in \mathcal{I},\ \forall j \in \mathcal{J} \tag{3.18}$$

$$v_{ij} \geqslant 0 \qquad \forall i \in \mathcal{I},\ \forall j \in \mathcal{J} \tag{3.19}$$

$$w_{i,j} \geqslant 0 \qquad \forall i \in \mathcal{I},\ \forall j \in \mathcal{J} \tag{3.20}$$

$$n_j \geqslant 0 \qquad \forall j \in \mathcal{J} \tag{3.21}$$

$$y_i \in \{0,1\} \qquad \forall i \in \mathcal{I} \tag{3.22}$$

$$u_{ij} \in \{0,1\} \qquad \forall i \in \mathcal{I},\ \forall j \in \mathcal{J} \tag{3.23}$$

$$z_{ij} \in \{0,1\} \qquad \forall i \in \mathcal{I},\ \forall j \in \mathcal{J} \tag{3.24}$$

3.4 Market Impacts under Adaptation and Non-Adaptation

Following the experimental design described above, we now analyze the effects of larger shares of intermittent renewables in a system with 77 GW of peak demand and geographically diversified renewable resources, which enable renewable production to be scaled for a given renewable capacity based on the historical capacity factor of wind and solar resources. In the adapted case, IMRES assumes a greenfield situation and decides on the individual generating units required to satisfy the demand at minimum cost. In the non-adapted case, IMRES assumes a capacity optimized for zero renewables but operated with a non-zero renewable capacity.

3.4.1 Optimal Capacity Mix

Figure 3-2 shows that, with adaptation, more renewable generation shifts the optimal generation mix towards having more flexible plants (CCGTs and OCGTs), displacing technologies traditionally considered as baseload. Therefore, a fraction of the blend of renewables and flexible technologies operate now as 'baseload'. This is primarily driven by how the shape of the net load duration curve changes as renewable capacity grows, as well as by the technical limitations of baseload technologies in the face of a more demanding operational regime. These two effects are consistent with those found in the MIT Future of Natural Gas Study (MIT, 2011). Additionally, coal plants do not appear in the optimal capacity mix as their startup costs are high comparatively to those of CCGTs, which can complement better the variability of renewable generation.

3.4.2 Capacity Factor and Renewable Curtailment

Without adaptation (solid lines in Figure 3-3), the average capacity factor (or level of utilization of the power plants) varies significantly across technologies: the capacity factor of nuclear power plants declines slightly with larger renewable penetration values, but it always stays at a level above 95 %; the capacity factor of coal plants declines sharply from 90% with

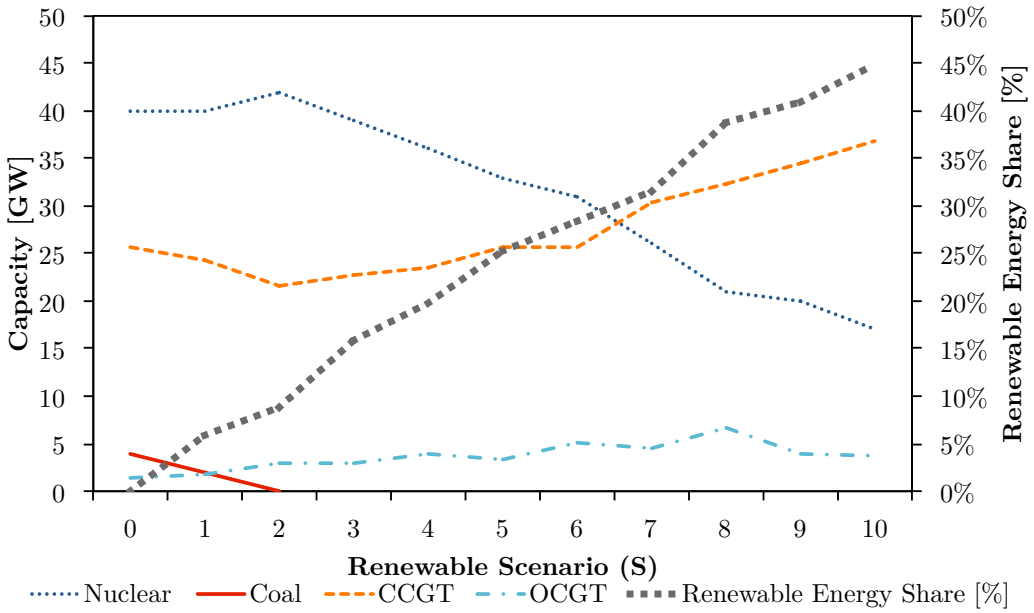


Figure 3-2: Evolution of the thermal capacity installed with greater renewable capacity installed

zero renewables to 30% in $S = 10$; the capacity factor of CCGTs and OCGTs also declines steadily from a level of 50% to 10% in the case of CCGTs, and from 8% to 4% in the case of OCGTs.

With adaptation (dashed lines in Figure 3-3), nuclear plants maintain a capacity factor level over 99% in most cases, coal plants are displaced, and CCGTs and OCGTs maintain a stable level above 40% and 8% respectively. Therefore, with adaptation all plants exhibit a higher capacity factor (i.e., a better utilization) than in the non-adapted system, as was to be expected. Note that with adaptation coal disappears from the optimal mix with renewable penetration levels greater than those in $S = 1$.

Similarly, renewable capacity is better utilized in the adapted case, and both solar and wind power exhibit larger capacity factors in the adapted case than in the non-adapted case. This can be better illustrated by looking specifically to the levels of renewable curtailment produced in the system (Figure 3-4). It can be seen how even though renewable curtailment grows as renewable capacity increases, the growth pace in a system that is not adapted is much quicker than in an adapted system. In this last case, renewable curtailment does not

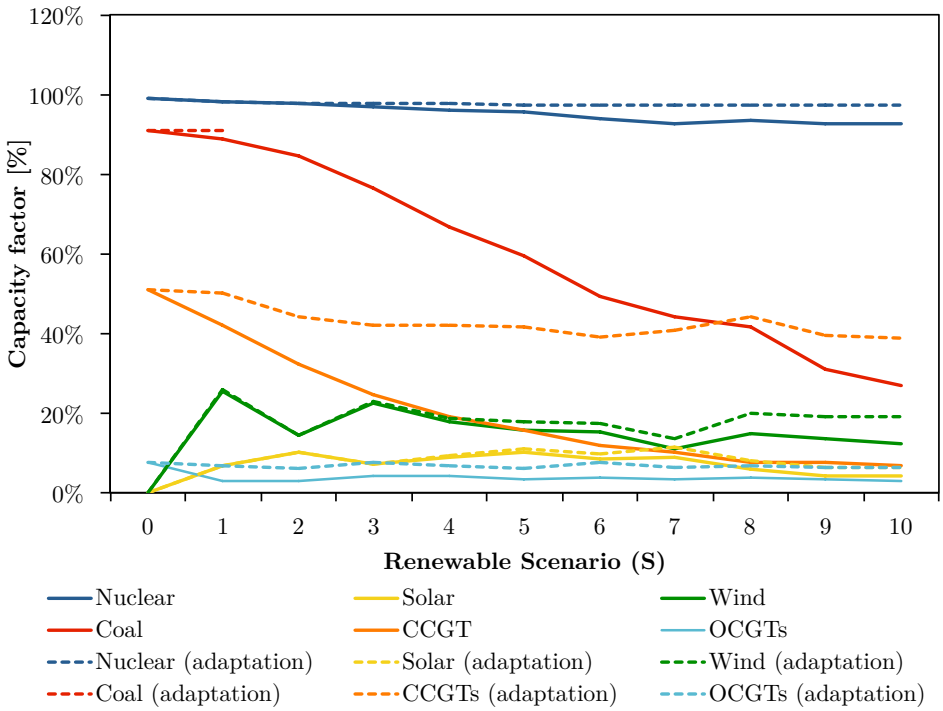


Figure 3-3: Capacity factor by technology under adaptation (dashed line) and non-adaptation (solid line).

go beyond 15% for the case with the largest renewable capacity in place.

3.4.3 Energy Contribution and Carbon Emissions

The technology-specific energy contribution (also known as the 'electricity mix'), reflects the participation of each technology in supplying the total electricity demand (Figure 3-5). It can be observed that without adaptation the participation of nuclear power in the electricity mix remains stable at about 68%, while the participation of CCGTs declines sharply from 23% to a mere 5%. Remember that without adaptation the capacity mix used is optimal for a situation with zero renewables, characterized by a homogeneous demand pattern with a high baseload that is covered by nuclear and coal. As more renewable capacity is unexpectedly added to the system, the generation mix that was originally optimized for zero renewables is operated to supply a lower and more variable net load. With adaptation, the opposite occurs: nuclear power contributes less to the total electricity mix with larger renewable

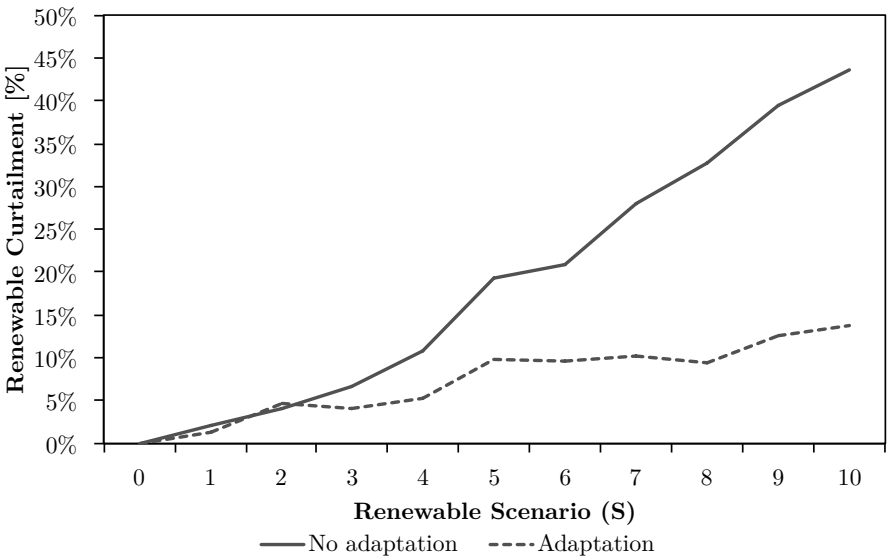


Figure 3-4: Renewable curtailment, expressed as % of the total renewable energy produced, under adaptation (dashed line) and non-adaptation (solid line).

shares, decreasing from 70% to 30%, while the energy contribution of CCGTs grows from 20% to 30%.

The total amount of carbon emissions is very much linked to the renewable penetration in each scenario and to the thermal technologies accompanying renewables in the energy mix. In the non-adapted case, emissions drop steadily, achieving 50% less emissions in $S = 10$ than in $S = 0$. Ultimately, as renewables are pushed into the system, they displace coal and gas-based technologies, coexisting only with nuclear power and balancing demand and generation through curtailment. With adaptation, total carbon emissions drop faster than in the non-adapted case, until they reach a floor of about 60% the emissions with no-renewables. This floor is explained because as the system adapts and nuclear power is displaced by CCGTs and OCGTs, the emissions associated with these gas-fueled technologies are non-zero, and the total emissions of the system will be those of a mix composed of gas-fueled plants and renewables (Figure 3-6).

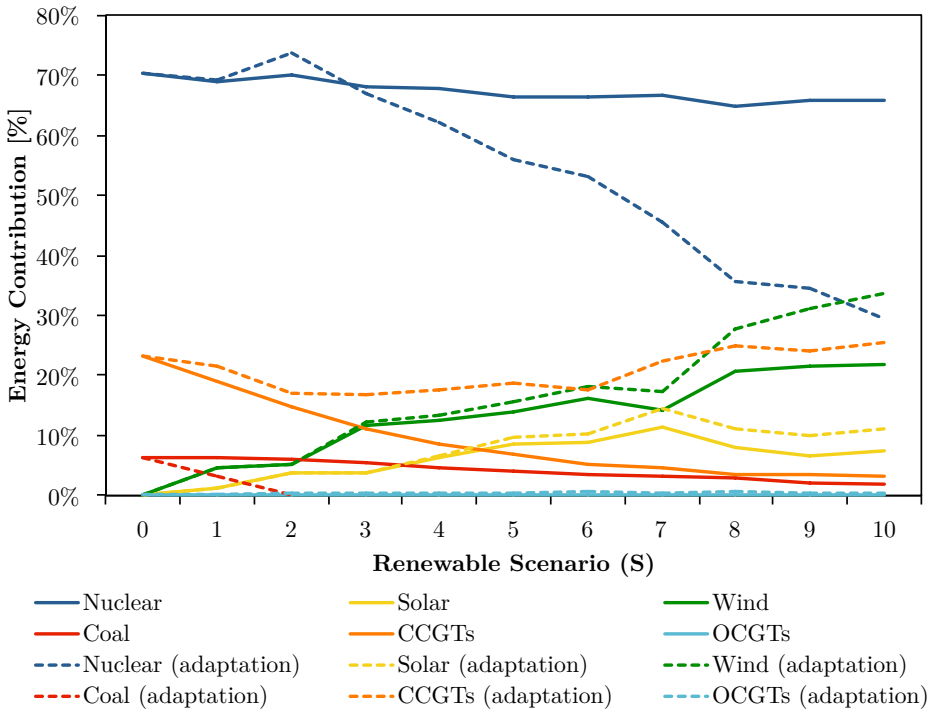


Figure 3-5: Energy contribution, expressed as % of total demand, by technology under adaptation (dashed line) and non-adaptation (solid line).

3.4.4 Total Startups

Results on the total number of startups (Figure 3-7) confirm the initial hypothesis in this thesis that the total startup cost in the system increases in the presence of intermittent renewables. With adaptation, we observe that the total number of startups performed by CCGTs and OCGTs grows. Without adaptation, we observe that the total number of startups of CCGTs and OCGTs remains relatively constant, while the number of startups performed by coal power plants increases, as they bear most of the extra cycling in the system.

The average number of startups by power plant (Figure 3-8) shows that with adaptation the number of startups performed by CCGTs grows with a larger renewable capacity, the number of startups by coal plants falls as coal plants are displaced by other technologies, and the startups by OCGTs do not follow any clear pattern. Without adaptation, the number of startups by CCGTs stays stable with larger shares of renewables, the number of startups

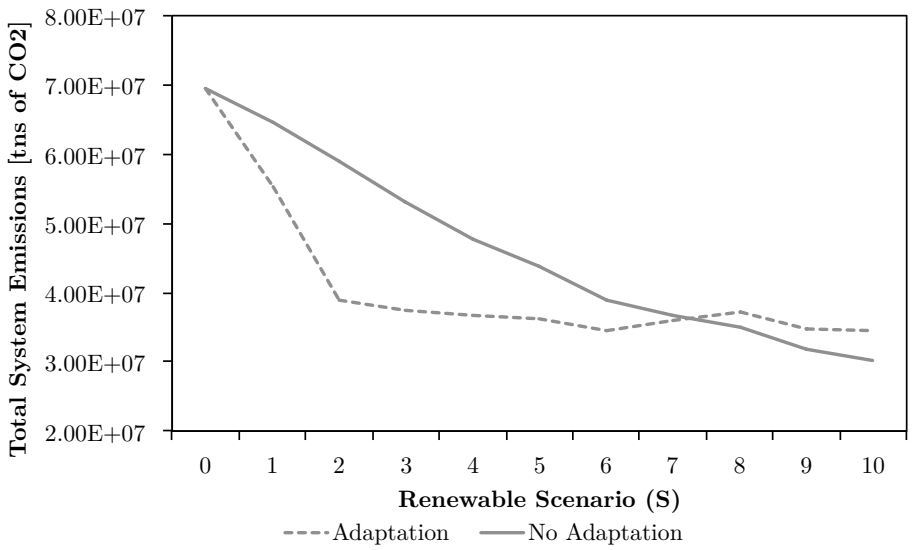


Figure 3-6: Total carbon emissions under adaptation (dashed line) and non-adaptation (solid line).

by coal plants increases, and the number of startups by OCGTs do not follow any particular pattern. Therefore, we these results show that larger shares of renewables induce a more intense cycling regime to some units in the system: with adaptation most of the extra cycling in the system is borne by CCGTs while, without adaptation, the extra cycling is borne by coal power plants.

3.4.5 One-Week Commitment Example

Figure 3-9 represents the wholesale price profile and the unit commitment of the resulting capacity mix in three different situations: no renewables in the system (Panel A), a capacity mix adapted for zero renewables but operating with 80 GW of wind and 80 GW of solar PV (Panel B), and a capacity mix adapted for 80 GW of wind and 80 GW of solar PV (Panel C). The situation in Panel B reflects a severe case of non-adaptation, which would be unlikely in reality, but as an extreme case it illustrates the effects more clearly.

Panel A shows that, without renewables, most of the baseload is provided by nuclear power, while coal plants operate as mid merit, reducing their output to the minimum stable

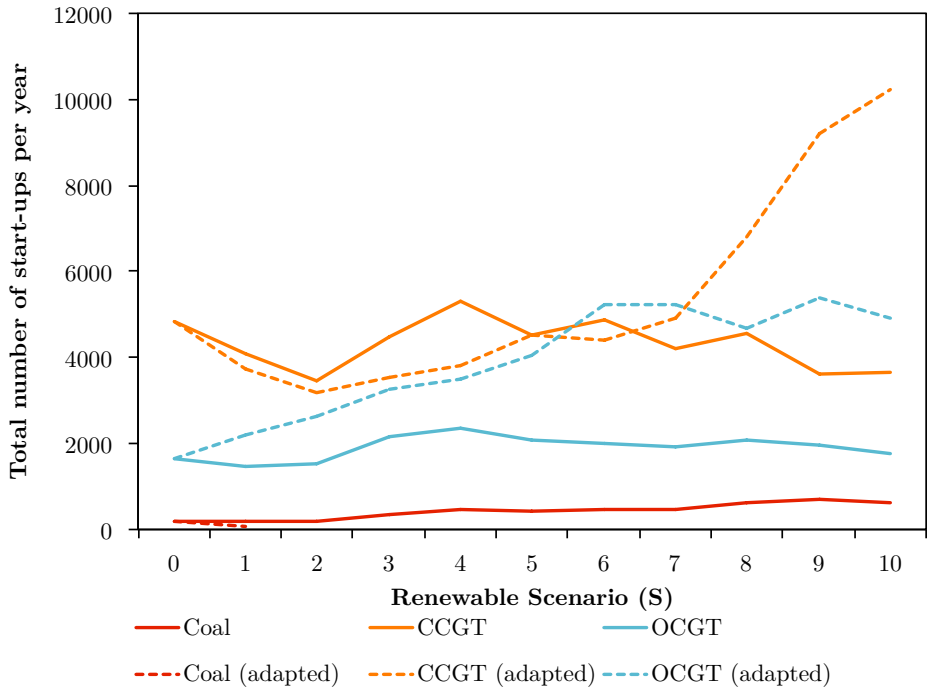


Figure 3-7: Total number of startups per year, by technology under adaptation (dashed line) and non-adaptation (solid line).

level when demand is low. CCGTs provide most of the power that is not baseload, adapting to the daily peaks and reducing their output or shutting down during the valleys. The wholesale prices derived from this situation follow the same pattern as demand, producing even one price spike during one hour where demand is very high and it is more economic to the system to have non-served energy than starting a new group. Yet, this small amount of non-served energy triggers the VOLL as wholesale price.

In Panel B the capacity installed is the same as the one in Panel A, but the capacity mix operates with a large amount of renewable capacity. We observe that CCGTs and OCGTs are almost completely displaced by renewables, as these are the two most expensive technologies in the system, and that coal plants have to cycle more, as it was indicated in the startup analysis. Wholesale prices are significantly lower in this case, because renewables have been added on top of an already adapted capacity mix, which has produced an excess capacity situation. In this case, technologies with the highest marginal costs are pushed

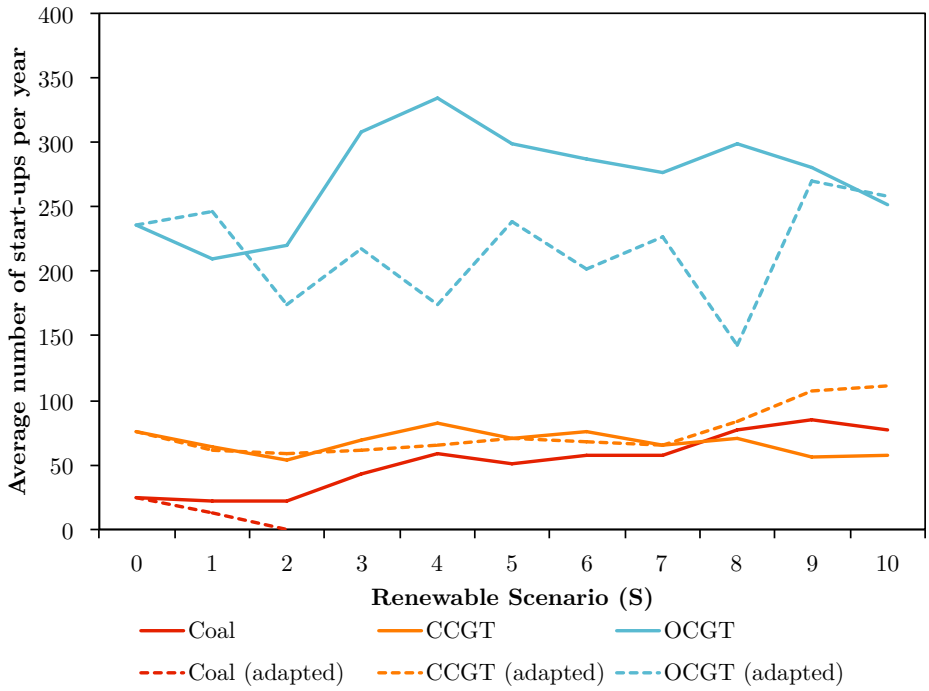


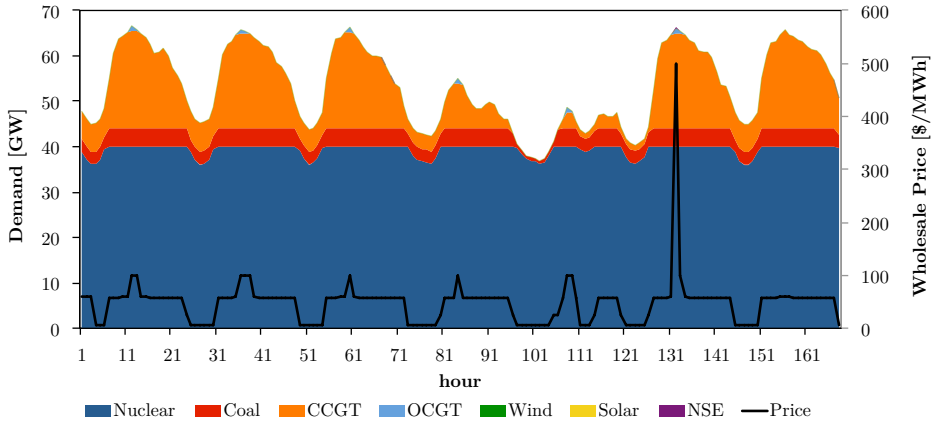
Figure 3-8: Average number of startups per year, by technology under adaptation (dashed line) and non-adaptation (solid line).

out of merit by renewables and the technologies setting the system marginal price have a lower variable cost. In addition, hours with very high renewable output hit the minimum output constraint of nuclear power, and the only degree of freedom left to meet the demand balance equation is the curtailment of renewable resource. The panel illustrates how during most hours there is renewable curtailment, with nuclear is producing at its minimum output level. Situations with renewable curtailment produce a zero marginal price in the system, and are responsible together with the excess of capacity of the low wholesale prices observed. Conversely, hours with small peaks in nuclear production indicate hours where there is no curtailment, and where it is economical to increase nuclear output above its minimum level.

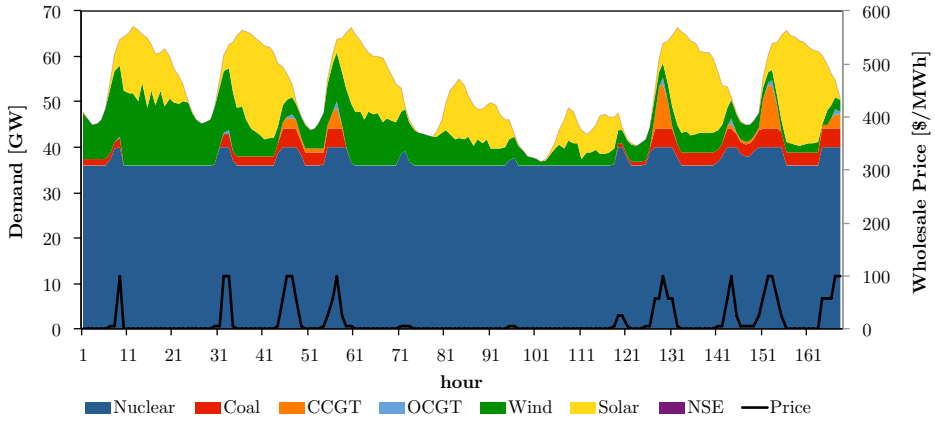
Lastly, Panel C shows the operation of a capacity mix that has been optimized for the same renewable capacity introduced in the case shown in Panel B. In this capacity mix, the amount of nuclear capacity is reduced compared to the two previous cases, and the number of CCGTs in the system is larger as they are more flexible and they can adapt to renewables'

variability at a lower cost than other technologies. This panel illustrates how now CCGTs and OCGTs complement wind and solar output to supply part of the baseload, and adapt to the joint variability of demand and renewables. Although there is a significant renewable deployment in the system, the resulting price series shows that they are very similar to those found in the case with zero renewables, evidencing that the adaptation state of the system to renewable penetration is one of the main drivers of the wholesale prices in the system.

Panel A: unit commitment for $S = 0$, adapted capacity mix, 0% renewable penetration.



Panel B: unit commitment for $S = 8$, non-adapted capacity mix, 16.75% renewable penetration.



Panel C: unit commitment for $S = 8$, adapted capacity mix, 36.93% renewable penetration.

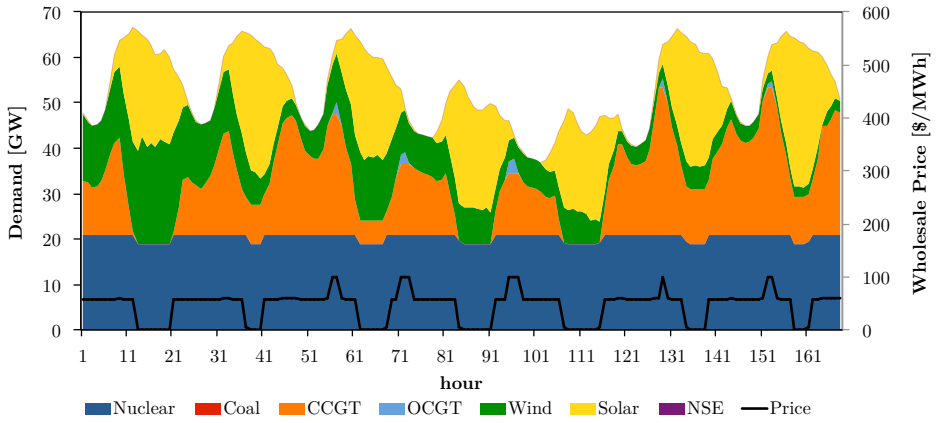


Figure 3-9: One week unit commitment and wholesale price profile in three different cases: $S = 0$, $S = 8$ without adaptation, and $S = 8$ with adaptation.

3.5 Bidding Rule Implementation

Bidding rules govern the structure of the bid functions submitted by the generators to electricity day-ahead auctions (i.e. the economic and technical conditions that are acceptable by generators in order to generate), influencing the market clearing price used to compensate generators for the energy delivered. *Simple bids* are price-quantity pairs that a generating unit is willing to accept as a fair compensation for producing electricity. Simple bids, however, do not represent explicitly the nonconvex components of cost, such as startup costs and no-load costs. As a consequence, generators are forced to internalize those cost items in the simple bid or, plants bid their actual marginal cost of energy and the market operator applies other mechanisms that allow plants to recover other costs. Alternatively, *complex bids* (also referred to as complimentary bids) are mechanisms that allow producers to explicitly reflect the nonconvex components of their cost structure in the bid.

Once all the bids have been received, the market operator centrally decides upon the commitment state and quantity dispatched of all the units participating in the auction, typically running a security-constrained unit commitment in the case of complex bids, and a straightforward matching of supply and demand curves in the case of simple bids. However, for the purpose of determining the generation mix and the minimum cost unit commitment and dispatch, it will be assumed that these decisions are centralized and the resulting generation mix, unit commitment and dispatch are optimal in the sense of minimizing costs. In principle, it is conceptually legitimate to assume that a perfectly competitive market and a centrally-planned system will result in the same generation expansion and operation, when all cost information is available to the decision maker[4](Perez-Arriaga and Meseguer, 1997). This section examines why the choice of bidding rules may create a discrepancy and its quantitative relevance. Therefore, in our experiments, we chose the generation mix and the operation of generating plants to be identical for all the bidding rules studied, differing only in the way that wholesale prices are determined. In reality, wholesale prices in the simple

[4]This is not strictly the case with simple bids, since the unit commitment and dispatch are not based on a explicit accounting of the start-up costs.

bids case also change the way that plants are scheduled and, ultimately, the capacity mix that the system would have in equilibrium. Nevertheless, we assume that the commitment states and capacity are the same across all cases, to observe more clearly what is the direct impact of bidding rules on prices and profits.

The bidding rules studied (with the exception of the simple bids case) are based on the dual variable μ_j of the demand balance equation (3.25). For problems with binary variables –such as IMRES–, the solver CPLEX uses a branch and cut algorithm, solving a series of LP subproblems where all binary variables are fixed (GAMS, 2014). Therefore, the duals obtained are those of an LP subproblem where commitment, startup and building variables are fixed.

$$\sum_{i \in \mathcal{I}} x_{ij} + n_j = D_j \qquad \perp \mu_j \text{ free} \qquad \forall j \in \mathcal{J} \tag{3.25}$$

3.5.1 Do Nothing

The "do nothing" case represents a naive approach to bidding, ignoring the nonconvex costs. The wholesale market price is $\rho_j^A = \mu_j,\ \forall j \in \mathcal{J}$, and the profitability of units is the difference between the remuneration obtained from energy sales and the sum of fixed costs and operating costs (3.26).

$$\pi_i^A = -C_i^{FIX} + \sum_{j \in \mathcal{J}} \left[\left(\rho_j^A - C_i^{VAR} \right) x_{ij} - C_i^{STUP} z_{ij} \right] \qquad \forall i \in \mathcal{T} \tag{3.26}$$

We use the results obtained from this strategy as comparison to quantify the effect produced by the subsequent bidding rules analyzed. This case is a simple theoretical situation used to show that if bidding rules do not account somehow for nonconvexities, operating costs cannot be recovered.

3.5.2 Side Payment

Some power systems like PJM and NY-ISO include a revenue sufficiency guarantee condition to ensure that power plants scheduled in the day-ahead and real-time markets recover their

daily offer amounts. In this case, the price is equal to the marginal cost of energy: $\rho_j^B = \mu_j,\ \forall j \in \mathcal{J}$, and there is a side payment that guarantees that plants recover their operating costs.

Side payments are allocated by looking at the difference between remuneration and operating costs on a daily basis. If we define an index $d \in \mathcal{D} = \left[1, ..., \left\lfloor \frac{N}{24} \right\rfloor\right]$ for the days included in the time series, the operating costs of power plant i incurred in day d, c_{id}, are given by the following expression:

$$c_{id} = \sum_{j \in \mathcal{J}^d} \left(C_i^{VAR} x_{ij} + C_i^{STUP} z_{ij} \right) \qquad \forall i \in \mathcal{T}, \forall d \in \mathcal{D} \tag{3.27}$$

where $\mathcal{J}^d$ denotes the subset of hours in day $d \in \mathcal{D}$:

$$\mathcal{J}^d = \left\{ j \middle| d = \left\lfloor \frac{j}{24} \right\rfloor ; j \in [1, ..., N] \right\} \tag{3.28}$$

We then assign a side payment equal to the difference between operating costs and energy remuneration, to plants where this difference is positive (3.29b). This mechanism ensures that plants are not overpaid when operating costs have already been recovered with energy sales (3.29a).

$$s_{id} = \begin{cases} 0, & \text{if } \sum_{j \in \mathcal{J}^d} \rho_j^B x_{ij} \geqslant c_{id} \quad (3.29a) \\ \sum_{j \in \mathcal{J}^d} c_{id} - \rho_j^B x_{ij}, & \text{if } \sum_{j \in \mathcal{J}^d} \rho_j^B x_{ij} < c_{id} \quad (3.29b) \end{cases}$$

The yearly profit of individual plants is equal to the difference between remuneration (including the side payments received) and cost:

$$\pi_i^B = -C_i^{FIX} + \sum_{j \in \mathcal{J}} \rho_j^B x_{ij} + \sum_{d \in \mathcal{D}} (s_{id} - c_{id}) \qquad \forall i \in \mathcal{T} \tag{3.30}$$

3.5.3 Price Uplift

The Irish market is another example of a centrally dispatched system where generators bid their actual short-run marginal cost. However, the Irish bidding code of practice (AIP, 2007) contemplates an addition to the price of an hourly uplift, denoted here as η_j, that is calculated on a daily basis to guarantee that all plants recover their operating costs. The resulting hourly system marginal price then becomes the sum of the marginal cost of energy and the hourly uplift: $\rho_j^C(\eta_j) = \mu_j + \eta_j, \ \forall j \in \mathcal{J}$.

The calculation of the uplift is performed ex-post of the daily unit commitment, for all $d \in \mathcal{D}$, fixing the variables resulting from the unit commitment and minimizing the cost of the uplift, and is subject to the constraint that all online power plants recover their operating costs (3.31).

$$\begin{aligned} &\min_{\boldsymbol{\eta}} \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}^d} \eta_j x_{ij} \\ &\text{s.t.:} \sum_{j \in \mathcal{J}^d} \left[(\rho_j^C(\eta_j) - C_i^{VAR}) x_{ij} - C_i^{STUP} z_{ij} \right] \geqslant 0 \qquad \forall i \in \mathcal{T} \\ &\qquad \eta_j \geqslant 0 \quad \forall j \in \mathcal{J}^d \end{aligned} \tag{3.31}$$

Similar to the side payment rule, price uplifts compensate for nonconvex cost items that are not covered by the marginal cost of energy (μ_j). However, in contrast to the side payment, the hourly uplift will affect the remuneration of all the on-line units. Another characteristic of this approach is the frequent presence of 'price spikes' as a consequence of adding the uplift on top of the price.

The resulting profit of individual plants is the difference between energy revenues obtained with the system marginal price and the sum of all fixed and variable costs:

$$\pi_i^C = -C_i^{FIX} + \sum_{j \in \mathcal{J}} \left[\left(\rho_j^C(\eta_j) - C_i^{VAR} \right) x_{ij} - C_i^{STUP} z_{ij} \right] \qquad \forall i \in \mathcal{T} \tag{3.32}$$

3.5.4 Simple Bids

In the absence of a separate remuneration mechanism like the ones used in PJM, NY-ISO or Ireland, auctions based on simple bids remunerate plants based solely on the wholesale price, forcing plants to internalize in the simple bid all nonconvex costs incurred by the plant. We implement this procedure by adding to the power plant's variable cost the average of the startup costs incurred by the plant over the course of one day (3.33) evenly distributed over the hours of operation of the plant. Note that, as indicated earlier, the plant dispatch resulting from this simple bidding does not have to coincide necessarily with the minimum cost unit commitment, where nonconvex operation costs are considered explicitly.

$$\rho_j^D = \max_{i \in \mathcal{I}} \left\{ \left(C_i^{VAR} + \frac{\sum_{j \in \mathcal{J}^d} C_i^{STUP} z_{ij}}{\sum_{j \in \mathcal{J}^d} x_{ij}} \right) u_{ij} \right\} \qquad \forall j \in \mathcal{J} \tag{3.33}$$

The annual profit of individual plants is the difference between the revenues from the energy produced and the sum of fixed and operating costs.

$$\pi_i^D = -C_i^{FIX} + \sum_{j \in \mathcal{J}} \left[\left(\rho_j^D - C_i^{VAR} \right) x_{ij} - C_i^{STUP} z_{ij} \right] \qquad \forall i \in \mathcal{T} \tag{3.34}$$

3.6 Market Equilibrium Heuristic Method

With adaptation, units expecting to make negative profits under a certain bidding rule would not be installed in the first place. The opposite can occur if bidding rules over-compensate some technologies and there is a positive profit expectation for new entrants. In fact, in a market equilibrium solution where all power plants recover their costs, no new entrant would make positive profits.

We propose here a set of rules to iterate on the welfare-maximizing generation fleet under a certain bidding rule until we obtain an approximation to the market equilibrium solution. This heuristic method screens all the technologies and adds or removes one unit repeatedly, until the profits of every individual unit are positive and no new units can be added that would make a profit. The technology screening is performed in an increasing variable cost

fashion, because plants that are lower in the merit order (i.e., lower marginal costs) have a greater impact on the energy delivered by units that are higher in the merit order. The resulting capacity change for a given technology k, expressed in number of units, is denoted as $\Delta_k(\cdot)$. A detailed description of the algorithm for this heuristic is provided in Fig. 3-10.

Require: $\hat{\boldsymbol{y}}^l = \boldsymbol{y}^*$ for $l = 0$
  $\mathcal{T}^1 \subset \mathcal{I} \rightarrow$ subset of nuclear plants
  $\mathcal{T}^2 \subset \mathcal{I} \rightarrow$ subset of coal plants
  $\mathcal{T}^3 \subset \mathcal{I} \rightarrow$ subset of combined cycle gas plants
  $\mathcal{T}^4 \subset \mathcal{I} \rightarrow$ subset of open cycle gas plants
  assuming that: $C^{VAR}_{i\in\mathcal{T}^1} < C^{VAR}_{i\in\mathcal{T}^2} < C^{VAR}_{i\in\mathcal{T}^3} < C^{VAR}_{i\in\mathcal{T}^4}$
  $\pi_i(\cdot) \rightarrow$ profit function of plant i
  $\mathbb{I}_i \rightarrow$ vector of 0's with a 1 at the i-th position
  for $k = 1$ **to** 4 **do**
    loop
      if $\exists i \in \mathcal{T}^k \mid \hat{y}^l_i = 0$ **and** $\pi_i(\hat{\boldsymbol{y}}^l + \mathbb{I}_i) > 0$ **then**
        $\hat{\boldsymbol{y}}^{l+1} = \hat{\boldsymbol{y}}^l + \mathbb{I}_i$
        $l = l + 1$
      else if $\exists i \in \mathcal{T}^k \mid \hat{y}^l_i = 1$ **and** $\pi_i(\hat{\boldsymbol{y}}^l) < 0$ **then**
        $\hat{\boldsymbol{y}}^{l+1} = \hat{\boldsymbol{y}}^l - \mathbb{I}_i$
        $l = l + 1$
      else
        break loop
      end if
    end loop
  end for
  $\Delta_k(\hat{\boldsymbol{y}}^l) = \sum_{i\in\mathcal{T}^k} \hat{y}^l_i - \sum_{i\in\mathcal{T}^k} y^*_i$
Ensure: $\pi_i(\hat{\boldsymbol{y}}^l) \geqslant 0 \ \forall i \in \mathcal{T}$ **and** $\nexists i' | \pi_i(\hat{\boldsymbol{y}}^l + \mathbb{I}_{i'}) \geqslant 0 \ \forall i \in \mathcal{T}$

Figure 3-10: Heuristic method to obtain a market equilibrium approximation

We apply this heuristic method to bidding rules yielding average profit solutions outside the -10% and $+10\%$ bounds with the adapted capacity mix. The results are shown in the next section.

3.7 Market Impacts of Bidding Rules

3.7.1 Impact on Cost Recovery

Tables 3.1-3.4 display the profit and cost to consumers results from all the experiments conducted, disaggregated by adaptation case, bidding rule applied, and renewable scenario.

In the non-adapted case, there are technology differential effects that can be further disaggregated into three components: 1) a price reduction derived from a reduction of the net demand; 2) a price increase from having more frequent non-convex costs impacting marginal prices; and 3) an energy output reduction as some units will see part of their energy produced displaced by renewable generation. Here, technologies with high variable cost will be more likely to be displaced first than technologies with a low variable cost. The high-renewable scenarios in the non-adapted case do not purport to represent any real situation, especially if renewable penetration is very high (e.g., it would be hard to find a situation where 80 extra Gigawatts of wind and solar power are deployed without having any kind of response from investors in thermal generation). These results, however, try to provide a complete illustration of the full spectrum of non-adaptation levels, from a more likely low unexpected renewable deployment, to a less likely high unexpected renewable deployment.

In the adapted case, the average profit values in the 'do nothing' case are negative as a result of remuneration not reflecting the nonconvex parts of the cost and the lumpiness of the plants. If we compare these results with those obtained with the non-adapted system, we observe that the magnitude of the remuneration gap for all technologies caused by having a non-adapted system significantly outweighs the gap caused by not remunerating startup costs with the adapted system.

Table 3.1: Results for the 'Do Nothing' Bidding Rule: Average Profit by Technology and Scenario and Average Cost to Consumers by Scenario

A. 'DO NOTHING'

Adapted System, $\boldsymbol{y}^*(S)$

	$\bar{\pi}_k(\boldsymbol{y}^*(S))$ [p.u.]				$\overline{CC}(\boldsymbol{y}^*(S))$
Renewable Scenario	Nuclear	Coal	CCGT	OCGT	$[\$/MWh]$
$S = 0$	-0.005	-0.042	-0.049	-0.079	68.427
$S = 1$	-0.030	-0.058	-0.051	-0.057	67.308
$S = 2$	-0.083	N/A	-0.05	-0.032	63.625
$S = 3$	-0.076	N/A	-0.03	-0.016	62.996
$S = 4$	-0.086	N/A	-0.062	-0.065	61.535
$S = 5$	-0.08	N/A	-0.043	-0.051	60.158
$S = 6$	-0.102	N/A	-0.07	-0.099	58.811
$S = 7$	-0.053	N/A	-0.033	-0.048	62.063
$S = 8$	-0.039	N/A	-0.004	0.033	62.509
$S = 9$	-0.065	N/A	-0.035	-0.058	60.066
$S = 10$	-0.091	N/A	-0.062	-0.086	58.058

Non-Adapted System, $\boldsymbol{y}^0(S)$

	$\bar{\pi}_k(\boldsymbol{y}^0(S))$ [p.u]				$\overline{CC}(\boldsymbol{y}^0(S))$
Renewable Scenario	Nuclear	Coal	CCGT	OCGT	[$/MWh]
$S = 0$	-0.005	-0.042	-0.049	-0.079	68.427
$S = 1$	-0.281	-0.305	-0.444	-0.798	48.354
$S = 2$	-0.359	-0.373	-0.55	-0.837	43.440
$S = 3$	-0.431	-0.443	-0.617	-0.779	38.682
$S = 4$	-0.502	-0.504	-0.676	-0.773	34.137
$S = 5$	-0.575	-0.568	-0.726	-0.803	28.211
$S = 6$	-0.638	-0.627	-0.755	-0.799	24.903
$S = 7$	-0.703	-0.687	-0.78	-0.816	20.421
$S = 8$	-0.723	-0.707	-0.801	-0.795	18.735
$S = 9$	-0.772	-0.764	-0.832	-0.817	15.411
$S = 10$	-0.796	-0.779	-0.845	-0.826	13.827

Average Profit by Technology and Scenario, $\bar{\pi}_k(\cdot)$ [p.u. of total cost]; and Average Cost to Consumers by Scenario , $\overline{CC}(\cdot)$ [$/MWh]. Renewable Capacity in each Scenario (S) is equal to $10 \cdot S$ GW of Wind and $10 \cdot S$ GW of Solar PV. For Example, $S = 3$ refers to a scenario with 30 GW of wind capacity and 30 GW of solar PV capacity installed

Table 3.2: Results for the 'Side Payment' Bidding Rule: Average Profit by Technology and Scenario and Average Cost to Consumers by Scenario

B. 'SIDE PAYMENT'

Adapted System, $\boldsymbol{y}^*(S)$

		$\bar{\pi}_k(\boldsymbol{y}^*(S))$ [p.u.]			$\overline{CC}(\boldsymbol{y}^*(S))$
Renewable Scenario	Nuclear	Coal	CCGT	OCGT	[$/MWh]
$S = 0$	0.001	-0.038	-0.014	-0.041	69.495
$S = 1$	-0.030	-0.058	-0.016	-0.024	68.06
$S = 2$	-0.083	N/A	-0.012	0.001	64.336
$S = 3$	-0.072	N/A	0.014	0.027	63.986
$S = 4$	-0.083	N/A	-0.017	-0.032	62.503
$S = 5$	-0.069	N/A	-0.009	-0.013	61.394
$S = 6$	-0.095	N/A	-0.026	-0.056	60.056
$S = 7$	-0.056	N/A	0.000	-0.009	62.733
$S = 8$	-0.036	N/A	0.034	0.055	63.535
$S = 9$	-0.065	N/A	-0.007	-0.011	60.955
$S = 10$	-0.082	N/A	-0.032	-0.040	59.500

Non-Adapted System, $\boldsymbol{y}^0(S)$

		$\bar{\pi}_k(\boldsymbol{y}^0(S))$ [p.u]			$\overline{CC}(\boldsymbol{y}^0(S))$
Renewable Scenario	Nuclear	Coal	CCGT	OCGT	[$/MWh]
$S = 0$	0.001	-0.038	-0.014	-0.041	69.495
$S = 1$	-0.281	-0.304	-0.403	-0.732	49.176
$S = 2$	-0.359	-0.363	-0.500	-0.760	44.396
$S = 3$	-0.43	-0.423	-0.569	-0.682	39.619
$S = 4$	-0.496	-0.492	-0.624	-0.669	35.257
$S = 5$	-0.567	-0.557	-0.662	-0.706	29.477
$S = 6$	-0.623	-0.611	-0.709	-0.705	26.242
$S = 7$	-0.683	-0.672	-0.725	-0.723	22.060
$S = 8$	-0.700	-0.687	-0.763	-0.699	20.263
$S = 9$	-0.740	-0.746	-0.775	-0.722	17.549
$S = 10$	-0.759	-0.768	-0.791	-0.738	16.120

Average Profit by Technology and Scenario, $\bar{\pi}_k(\cdot)$ [p.u. of total cost]; and Average Cost to Consumers by Scenario , $\overline{CC}(\cdot)$ [$/MWh]. Renewable Capacity in each Scenario (S) is equal to $10 \cdot S$ GW of Wind and $10 \cdot S$ GW of Solar PV. For Example, $S = 3$ refers to a scenario with 30 GW of wind capacity and 30 GW of solar PV capacity installed

Table 3.3: Results for the 'Price Uplift' Bidding Rule: Average Profit by Technology and Scenario and Average Cost to Consumers by Scenario

C. 'PRICE UPLIFT'

Adapted System, $\boldsymbol{y}^*(S)$

	$\bar{\pi}_k(\boldsymbol{y}^*(S))$ [p.u.]				$\overline{CC}(\boldsymbol{y}^*(S))$
Renewable Scenario	Nuclear	Coal	CCGT	OCGT	$[\$/MWh]$
$S = 0$	0.068	0.027	0.014	-0.001	73.370
$S = 1$	0.039	0.009	0.007	0.004	71.926
$S = 2$	0.010	N/A	0.027	0.031	69.708
$S = 3$	0.010	N/A	0.04	0.065	68.705
$S = 4$	0.019	N/A	0.022	0.007	68.206
$S = 5$	0.025	N/A	0.023	0.026	66.62
$S = 6$	0.013	N/A	0.012	-0.005	65.858
$S = 7$	0.038	N/A	0.037	0.039	67.875
$S = 8$	0.06	N/A	0.078	0.087	68.620
$S = 9$	0.026	N/A	0.023	0.02	65.717
$S = 10$	-0.004	N/A	0.000	-0.011	63.580

Non-Adapted System, $\boldsymbol{y}^0(S)$

	$\bar{\pi}_k(\boldsymbol{y}^0(S))$ [p.u]				$\overline{CC}(\boldsymbol{y}^0(S))$
Renewable Scenario	Nuclear	Coal	CCGT	OCGT	[$/MWh]
$S = 0$	0.068	0.027	0.014	-0.001	73.370
$S = 1$	-0.202	-0.227	-0.372	-0.677	53.907
$S = 2$	-0.237	-0.254	-0.447	-0.642	51.895
$S = 3$	-0.323	-0.343	-0.546	-0.579	45.934
$S = 4$	-0.383	-0.392	-0.591	-0.554	42.127
$S = 5$	-0.468	-0.466	-0.648	-0.613	35.253
$S = 6$	-0.509	-0.524	-0.683	-0.558	33.529
$S = 7$	-0.587	-0.598	-0.711	-0.639	27.924
$S = 8$	-0.604	-0.614	-0.745	-0.601	26.281
$S = 9$	-0.65	-0.682	-0.763	-0.624	23.251
$S = 10$	-0.678	-0.698	-0.771	-0.639	21.467

Average Profit by Technology and Scenario, $\bar{\pi}_k(\cdot)$ [p.u. of total cost]; and Average Cost to Consumers by Scenario , $\overline{CC}(\cdot)$ [$/MWh]. Renewable Capacity in each Scenario (S) is equal to $10 \cdot S$ GW of Wind and $10 \cdot S$ GW of Solar PV. For Example, $S = 3$ refers to a scenario with 30 GW of wind capacity and 30 GW of solar PV capacity installed

Table 3.4: Results for the 'Simple Bids' Bidding Rule: Average Profit by Technology and Scenario and Average Cost to Consumers by Scenario

D. 'SIMPLE BIDS'

Adapted System, $\boldsymbol{y}^*(S)$

	$\overline{\pi}_k(\boldsymbol{y}^*(S))$ [p.u.]				$\overline{CC}(\boldsymbol{y}^*(S))$
Renewable Scenario	Nuclear	Coal	CCGT	OCGT	$[\$/MWh]$
$S = 0$	0.487	0.382	0.317	0.232	99.830
$S = 1$	0.460	0.363	0.299	0.236	98.288
$S = 2$	0.406	N/A	0.282	0.254	94.527
$S = 3$	0.359	N/A	0.273	0.256	90.640
$S = 4$	0.393	N/A	0.292	0.245	92.268
$S = 5$	0.320	N/A	0.215	0.249	86.327
$S = 6$	0.362	N/A	0.256	0.233	88.981
$S = 7$	0.337	N/A	0.244	0.239	87.809
$S = 8$	0.368	N/A	0.327	0.335	89.308
$S = 9$	0.354	N/A	0.239	0.248	86.966
$S = 10$	0.318	N/A	0.214	0.217	84.450

Non-Adapted System, $\boldsymbol{y}^0(S)$

	$\overline{\pi}_k(\boldsymbol{y}^0(S))$ [p.u]				$\overline{CC}(\boldsymbol{y}^0(S))$
Renewable Scenario	Nuclear	Coal	CCGT	OCGT	[$/MWh]
$S = 0$	0.487	0.382	0.317	0.232	99.830
$S = 1$	0.179	0.089	-0.186	-0.672	76.837
$S = 2$	0.113	0.029	-0.325	-0.674	72.785
$S = 3$	0.052	-0.045	-0.411	-0.612	69.568
$S = 4$	-0.025	-0.107	-0.479	-0.587	65.035
$S = 5$	-0.114	-0.186	-0.533	-0.641	58.529
$S = 6$	-0.212	-0.286	-0.596	-0.589	53.024
$S = 7$	-0.309	-0.369	-0.63 0	-0.670	47.014
$S = 8$	-0.333	-0.39 0	-0.666	-0.637	44.091
$S = 9$	-0.410	-0.487	-0.679	-0.672	40.109
$S = 10$	-0.470	-0.522	-0.693	-0.669	36.148

Average Profit by Technology and Scenario, $\overline{\pi}_k(\cdot)$ [p.u. of total cost]; and Average Cost to Consumers by Scenario , $\overline{CC}(\cdot)$ [$/MWh]. Renewable Capacity in each Scenario (S) is equal to $10 \cdot S$ GW of Wind and $10 \cdot S$ GW of Solar PV. For Example, $S = 3$ refers to a scenario with 30 GW of wind capacity and 30 GW of solar PV capacity installed

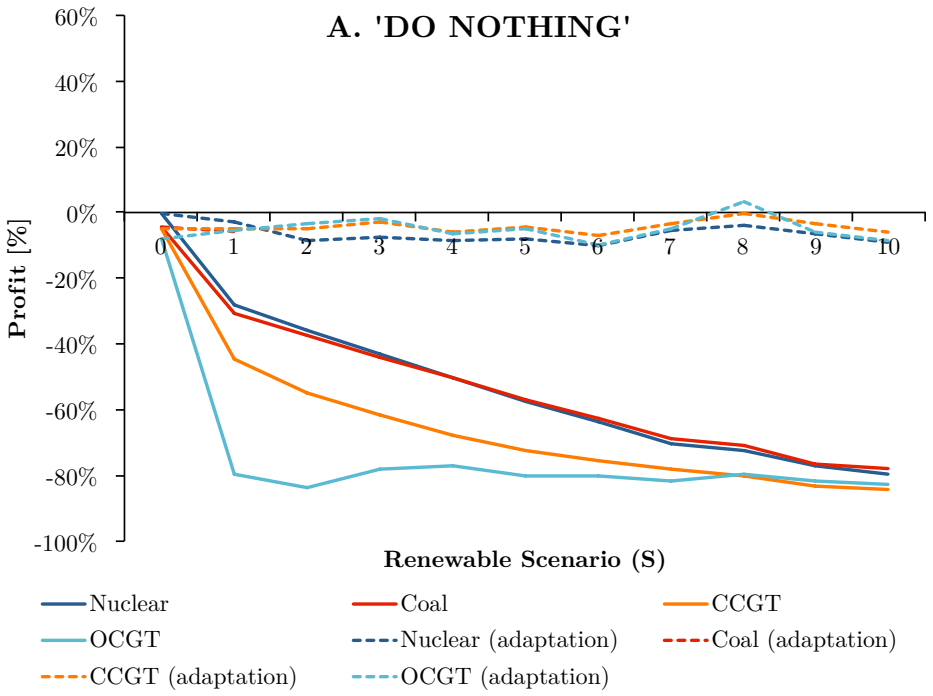


Figure 3-11: Profit across renewable scenarios in % of total plant cost, with the 'Do Nothing' bidding rule.

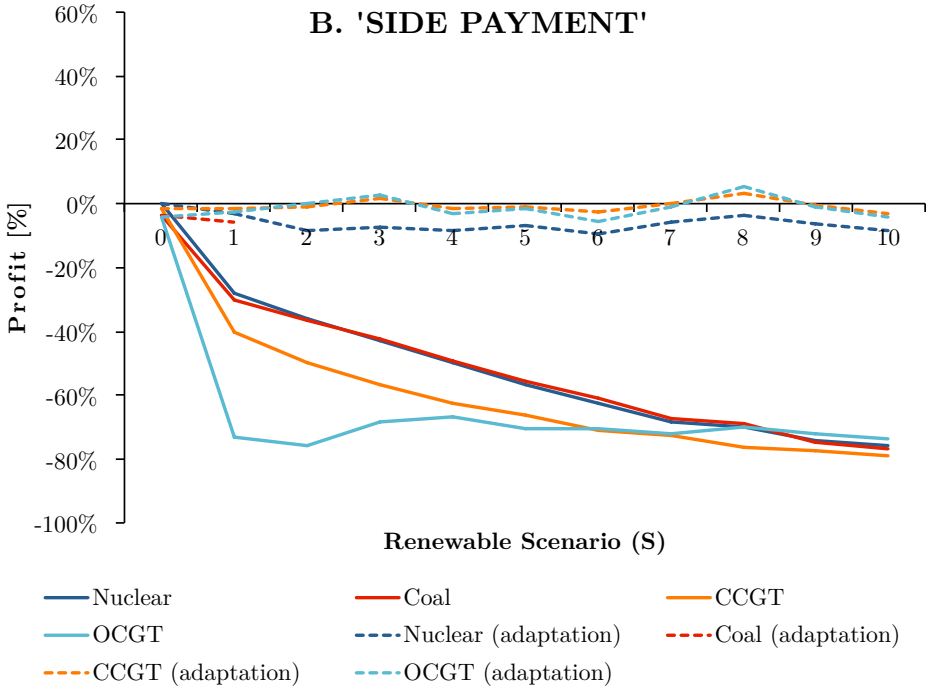


Figure 3-12: Profit across renewable scenarios in % of total plant cost, with the 'Side Payment' bidding rule.

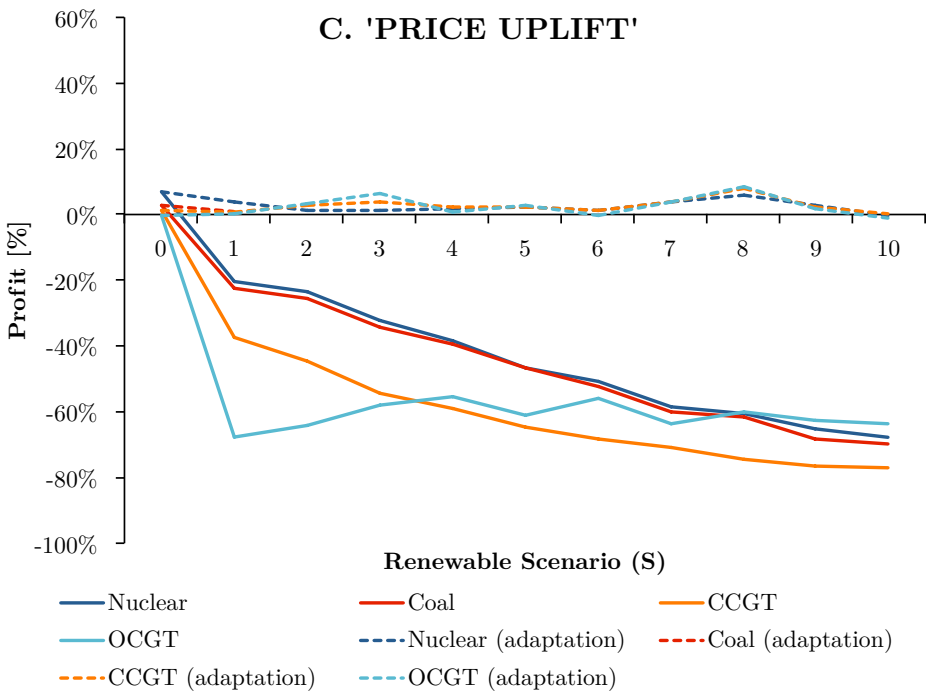


Figure 3-13: Profit across renewable scenarios in % of total plant cost, with the 'Price Uplift' bidding rule.

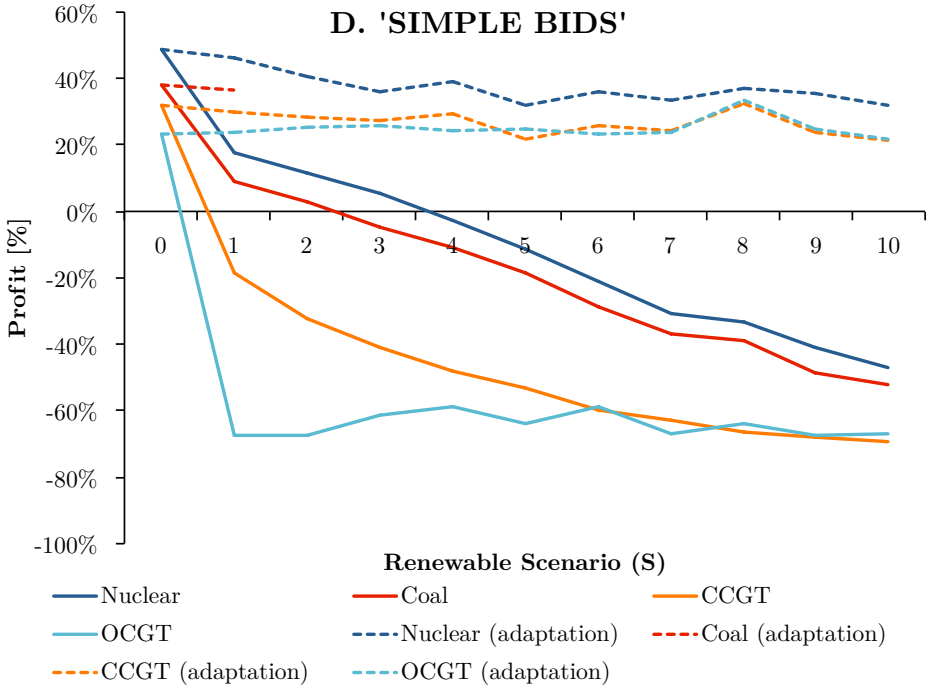


Figure 3-14: Profit across renewable scenarios in % of total plant cost, with the 'Simple Bids' bidding rule.

The 'side payment and the 'price uplift' bidding rules guarantee that under adaptation all plants recover operating costs, but the 'price uplift rule helps all technologies towards recovering their total costs as the non-convex costs reflected in the uplift are added to the marginal price. We illustrate in Figure 3-15 the average profit for each thermal technology considered and for the first three bidding rules analyzed.

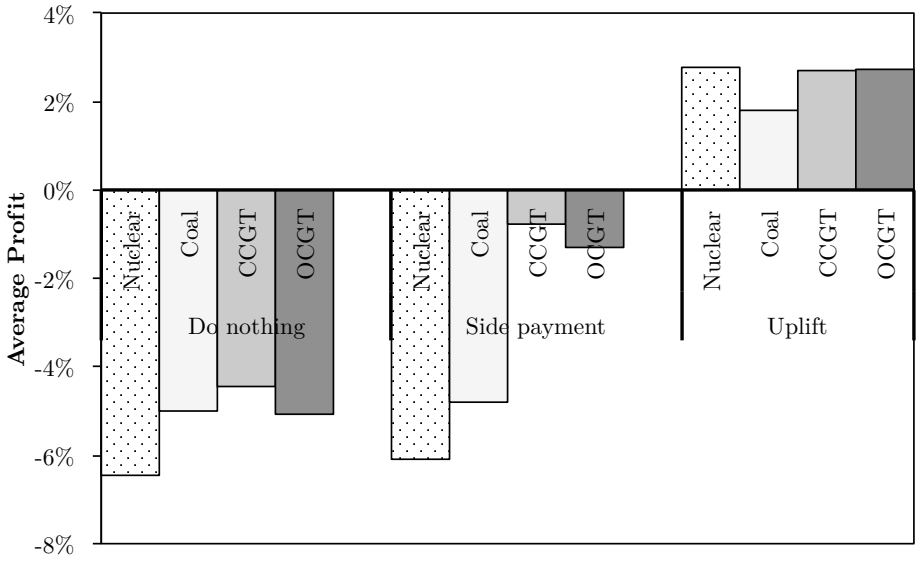


Figure 3-15: Average profit as a percentage of the total cost of the unit, across all renewable penetration scenarios with each of the bidding rules studied

Under adaptation, the 'simple bids' rule yields profits significantly greater than zero, leading to a situation that may attract new entrants to the market. Therefore, we apply the heuristic method described before to this case to obtain a market equilibrium approximation. We choose to consider only investments in additional CCGT or OCGT units to test the heuristic within a reduced search-space.

Table 3.5 shows in parenthesis the number of new plants added after applying the heuristic. $\overline{CC}$ results show that by letting plants reach a market equilibrium, we reduce the cost to consumers. Yet, they also reveal that, in equilibrium, the solution from using simple bids produces a situation of excess capacity where plants are still overcompensated.

Table 3.5: Average Profit by Technology and Scenario, $\bar{\pi}_k(\cdot)$ [p.u. of total cost]; Capacity Increase $\Delta_k(\cdot)$ (in parenthesis); and Average Cost to Consumers by Scenario , $\overline{CC}(\cdot)$ [\$/MWh] for bidding rule *D. Simple Bids* and renewable capacity scenarios (S).

Scenario	$\bar{\pi}_k(\hat{\boldsymbol{y}}^l(S))$ [p.u.], Nuclear	Coal	$\Delta_k(\hat{\boldsymbol{y}}^l(S))$ [units] CCGT	OCGT	$\overline{CC}(\hat{\boldsymbol{y}}^l(S))$ [\$/MWh]
D. Simple Bids					
$S = 0$	0.428	0.324	0.212	0.042	95.037
	(=)	(=)	(+1)	(+1)	
$S = 1$	0.406	0.31	0.208	0.1	93.991
	(=)	(=)	(+1)	(-1)	
$S = 2$	0.348	N/A	0.195	0.139	90.262
	(=)	(=)	(+2)	(-2)	
$S = 3$	0.282	N/A	0.143	0.074	85.251
	(=)	(=)	(+2)	(=)	
$S = 4$	0.321	N/A	0.176	0.071	86.932
	(=)	(=)	(+2)	(=)	
$S = 5$	0.249	N/A	0.131	0.094	81.663
	(=)	(=)	(+1)	(=)	
$S = 6$	0.315	N/A	0.172	0.117	85.643
	(=)	(=)	(+3)	(-2)	
$S = 7$	0.294	N/A	0.161	0.117	84.619
	(=)	(=)	(+1)	(+1)	
$S = 8$	0.279	N/A	0.181	0.059	83.057
	(=)	(=)	(+2)	(=)	
$S = 9$	0.29	N/A	0.131	0.089	82.789
	(=)	(=)	(+3)	(=)	
$S = 10$	0.284	N/A	0.136	0.089	82.461
	(=)	(=)	(+2)	(+1)	

3.7.2 Impact on the Cost to Consumers

Figure 3-16 shows the impact of renewable capacity on the average cost paid by consumers ($\overline{CC}$)[5]. We observe a significant divergence between the results from the adaptation and the non-adaptation cases, as in the non-adaptation case renewables are forced into a system that is already designed to cover the demand by itself, creating a situation of excess-capacity. Renewables thus displace the most expensive plants, reduce the marginal cost of energy, and reduce the number of hours with non-served energy, ultimately decreasing the wholesale price of electricity.

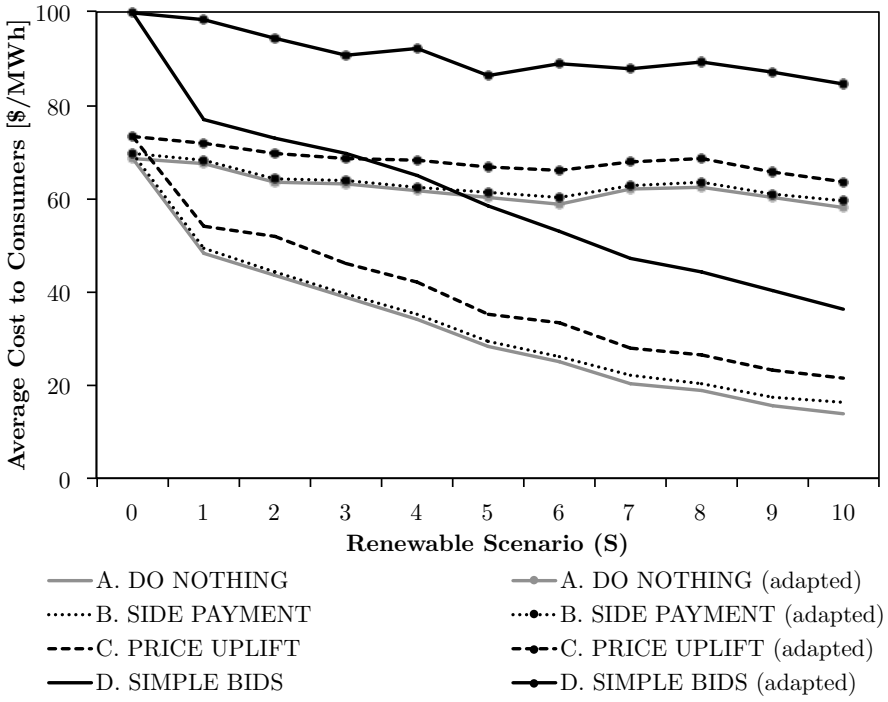


Figure 3-16: Average cost to consumers across all renewable penetration scenarios with each of the bidding rules studied

In the adapted cases, we observe that although $\overline{CC}$ tends to be lower with greater values of renewable capacity, the decline is less steep than in the non-adaptation cases. We also note that $\overline{CC}$ with the 'do-nothing' rule and with a 'side-payment' are both very similar,

[5]Note that the average cost to consumers calculated does not include the cost of any incentives provided to promote renewable investments. Also, the higher the penetration of renewables, the lower the market prices, therefore rendering investment in renewables less attractive.

and that bidding rules that add an uplift increase the cost to consumers by about 8% of the initial value. Simple bids impose a much greater burden on the consumer, on the order of 35% greater than the 'do-nothing' case.

3.8 Sensitivity Analysis

As discussed in the theoretical background, if the generation fleet is in equilibrium, and in the absence of common or internal constraints on capacity, then all plants must have a profit equal to zero under ideal conditions (see Introduction for a review of these conditions). Replicating this theoretical result with a real-sized system model, while accounting for the lumpiness associated with the limited number of plant sizes, and the most binding technical constraints, presents a significant challenge. The calibration of a complex model like the one used in this thesis has helped us learn many details that would have been hard to predict without the extensive experimentation in the analyses performed. This section presents a summary of some of the sensitivity analyses performed to calibrate the model and some interesting effects that were found.

3.8.1 Sensitivity to the VOLL

As the model uses an approximation of the $NLDC$, the cost associated with every hour contained in the approximation is scaled by the factor Θ to approximate the total operation cost in one year. Therefore, hours with non-served energy will be multiples of the factor Θ, introducing an extra discontinuity in the total cost objective function. This discontinuity reduces the accuracy of the representation of the peak demand, potentially overestimating the number of hours priced at the VOLL and the remuneration of units (e.g., if a full-year series has 3 hours with NSE, and the weeks selected contain one of those hours, the full-year approximation will have Θ hours with NSE, overestimating the remuneration of the units). High values of VOLL can amplify this effect, yielding profit results that are far from the theoretical result. Figure 3-17 shows profit values for three of the thermal generation

technologies considered under different values of lost load:

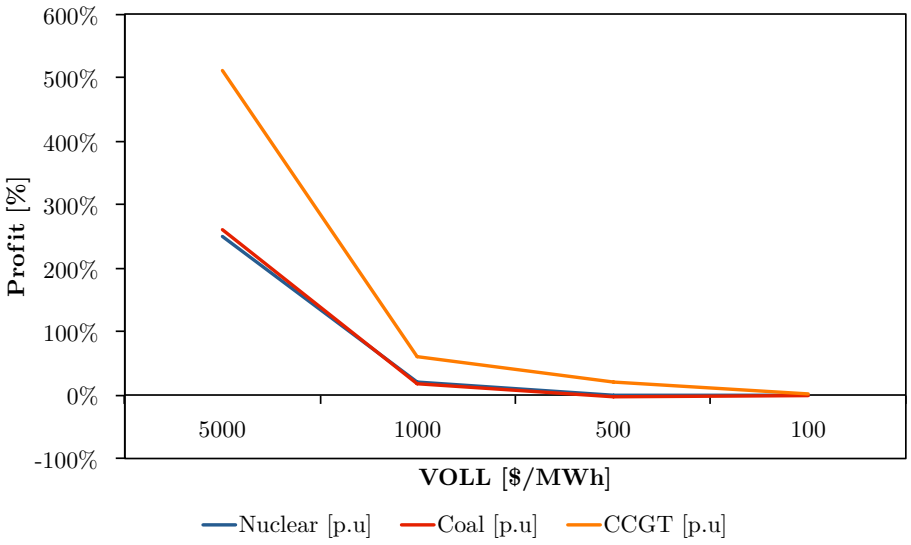


Figure 3-17: Sensitivity of technology-specific average profits to different values of non-served energy

One possible solution to overcome this effect and approximate equilibrium solutions closer to the zero-profit level, would be not scaling up hours with non-served energy in the cost function. However, this solution is not straightforward as other additional changes to how the approximation is built would be required, so that the objective function reflects the cost of 8,760 hours. Another possibility, is to use a low VOLL that mitigates the impact on profits of having more hours with NSE in the approximation than in the full-year series. This is the approach chosen in this thesis.

3.8.2 Sensitivity to the Solver Tolerance

The convergence rate of high-dimension model such as IMRES significantly increases as the relative gap tolerance decreases. At the same time, pricing models require a low gap tolerance to obtain prices that are as close as possible to those that would support equilibrium. The experiments in this thesis used a tolerance $\epsilon = 10^{-4}$ as a compromise solution that permits solving the model in a reasonable amount of time while producing prices that get plants'

profits close to zero. Solver tolerances greater than 10^{-4} solve IMRES more rapidly, but produce solutions with a larger duality gap and further from equilibrium. Conversely, solver tolerances below 10^{-4} find a solution to IMRES in a much longer time period at best, as sometimes the solver does not even find a solution.

3.8.3 Impact of hourly ramps

The effect of hourly ramps on the solution was tested by running IMRES in both the adaptation and non-adaptation cases, with and without these constraints (Table 3.6). The test did not show any significant difference between including hourly ramps or not. We observed that, with adaptation, the capacity installed with and without including ramp constraints was the same in both cases. The capacity factor and energy contribution of renewables was found to be slightly higher without ramp constraints, because when these constraints are binding the system decides to increase renewable curtailment. If the amount of energy curtailed had been higher than the one found in this test, the opportunity cost of the energy curtailed might have been sufficiently high for the model to decide to start up or shut down one plant.

There are at least two possible explanations for the almost negligible effect of ramps in this test: first, that the system counts with a sufficiently large number of units that allows the system to coordinate the production of many units upwards or downwards in order to withstand very steep ramps; and second, the steepest ramps might not necessarily occur at timeframes equal or greater than one hour, and might occur in timeframes of seconds or minutes. Therefore, ramps might play a more important role in smaller-sized systems where there are not enough plants to increase their output simultaneously, or within a smaller timeframe (for instance, a 5 minute timeframe) that is not reflected in this model. Considering ramps at a smaller time scale with a capacity expansion model like IMRES might increase the dimensionality challenge and other statistical methods or models that focus on smaller time frame decisions might account better for this type of phenomena.

Table 3.6: Ramp constraints sensitivity analysis. Total capacity, capacity factor and energy contribution results for the adapted and the non-adapted system, with and without hourly ramping constraints.

	Adapted System, $\boldsymbol{y}^*(S)$		Non-Adapted System, $\boldsymbol{y}^0(S)$	
	w/ ramp constraints	w/o ramp constraints	w/ ramp constraints	w/o ramp constraints
Renewable contribution [%]	44.3	44.5	30.7	31
Curtailment [%]	8.9	8.5	36.9	36.1
Non-served energy [%]	0.263	0.264	0.003	0.004
Average Price [$/MWh]	58.086	58.191	9.247	9.655
Capacity [GW]				
Nuclear	17	17	40	40
Coal	0	0	5.5	5.5
CCGT	36.4	36.4	25.6	25.6
OCGT	0	0	0	0
Capacity Factor [%]				
Nuclear	98.4	97.9	93	92.7
Solar	8.4	8.5	5.8	5.2
Wind	16.6	16.6	11.5	12.3
Coal	–	–	21.4	20.4
CCGT	40	40	2.8	2.7
OCGT	–	–	–	–
Energy Contribution [%]				
Nuclear	29.8	29.6	66	65.7
Solar	14.9	15.1	10.3	9.2
Wind	29.4	29.4	20.4	21.8
Coal	–	–	2.1	2
CCGT	25.9	25.9	1.3	1.2
OCGT	–	–	–	–

3.8.4 Impact of Startup Cost

As has been mentioned throughout this thesis, one of the most important nonconvexities found in the generation activity in real power systems is startup costs. This subsection tests the effect of the magnitude of this parameter on the final profits obtained by different technologies by running all the renewable scenarios with adaptation with all the bidding rules studied. The experiment compares the results obtained using the original values of startup costs ('high') (see Table A.3) with those obtained with a set of plants with startup costs equal to one third of the original values ('low').

Figures 3-18 and 3-19 present a comparison of the average and total number of startups per technology with low (dotted) and high (dashed) startup costs. Specifically, Figure 3-18 shows that if the startup cost of the individual plants is lowered to one third of the original value, the startup regime of CCGTs is increased and the startup regime of OCGTs is reduced on average. Figure 3-19 shows that the total number of startups increase when startup costs are lower (as it would be expected), and that the cycling burden is shifted from OCGTs in the high startup case to CCGTs in the low startup case. Additionally, the figure illustrates how the total system-wide cost for starting up is reduced, as the startup costs in the low case is one third of the high case, and the total number of startups in the 'low' case is less than three times the number of startups in the 'high' case.

Figures 3-20-3-23 present the profit results of this analysis. Note that since we are re-running IMRES with other values of startup costs, the values of the capacity variables as well as the variables defining the operation of the plants will change, in addition to the profits obtained by the plants.

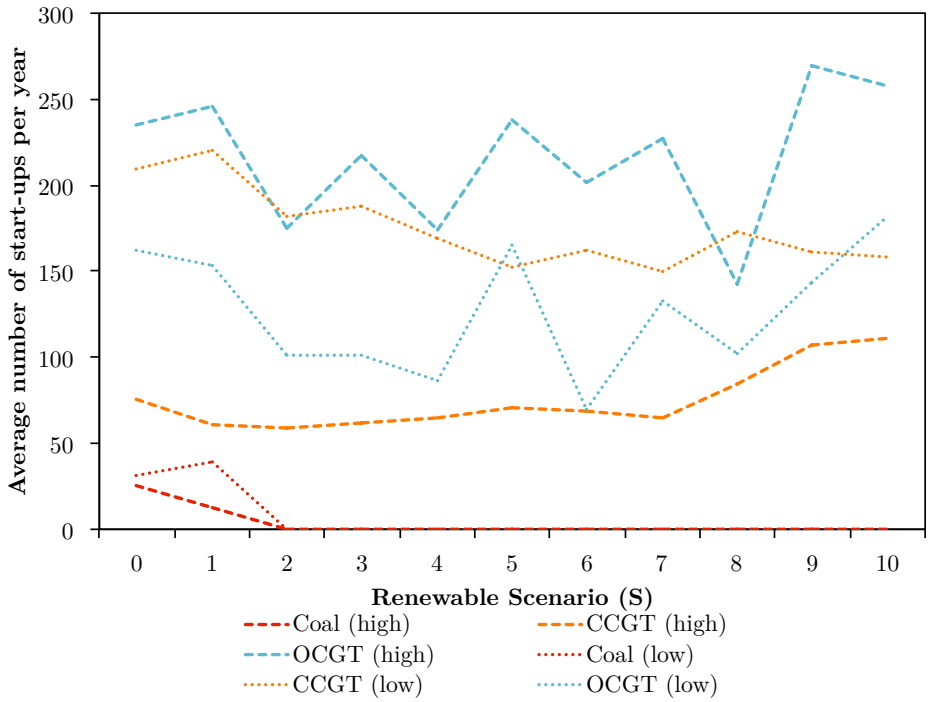


Figure 3-18: Sensitivity of the average number startups to the startup cost of the plants.

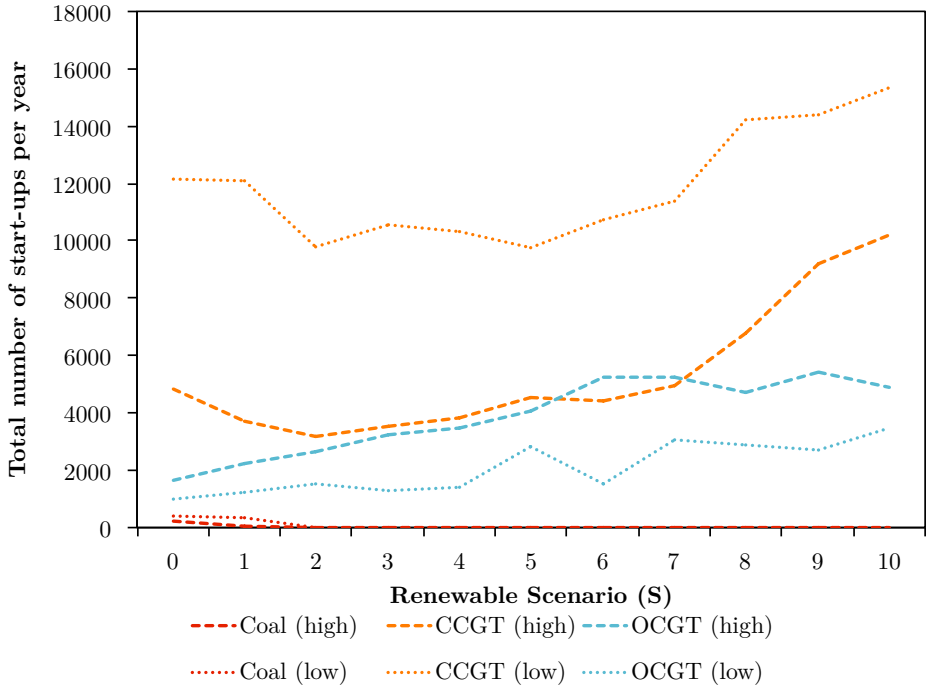


Figure 3-19: Sensitivity of the total number startups to the startup cost of the plants.

The 'do-nothing' case (Figure 3-20) shows that startup costs are responsible for a significant fraction of the divergence between the actual profits obtained by the plants, and the theoretical zero profit level obtained in equilibrium with a system where all costs are convex. In the absence of any bidding mechanism that allows recovering startup costs, we observe that technologies are closer to the theoretical result when startup costs are low, as the nonconvexities in this case are smaller. Yet, bidding rules that account for nonconvex costs are still necessary.

The 'side payment' bidding rule (Figure 3-21) shows a positive effect on the technologies that bear with the cycling in the system (CCGTs and OCGTs), and do not affect the total cost of other technologies that are below in the merit order (nuclear and coal). Remember that, in the experiments with high startup costs, coal gets displaced with penetration levels larger than $S = 1$. This effect is also found with low startup costs.

With a 'price uplift' (Figure 3-22), we observe that all technologies recover again total costs. The profits of CCGTs and technologies below them in the merit order increase relative to the high-startup case, while the profits obtained by OCGTs remain on the positive side but are now lower. This effect can be explained based on the changes in the startup regime from 'low' to 'high' startup costs: with a low startup cost, units in general will cycle more, and some of the cycling that was borne by OCGTs with high startup costs is now borne by CCGTs. Also, as units cycle more, in this case the average uplift is higher, increasing wholesale prices, which will be reflected on the profits of the technologies that cycle more, and on those of other technologies below them in the merit order. This is why OCGTs do not see the profit increase seen by other technologies with a lower variable cost.

Finally, the 'simple bids' case (Figure 3-23) shows that as the total cost of starting up is lower, the price increase that internalizes the average daily startup cost is also lower compared to the high startup cost case. Therefore, profits in the 'low' case are lower than in the 'high' startup cost situation.

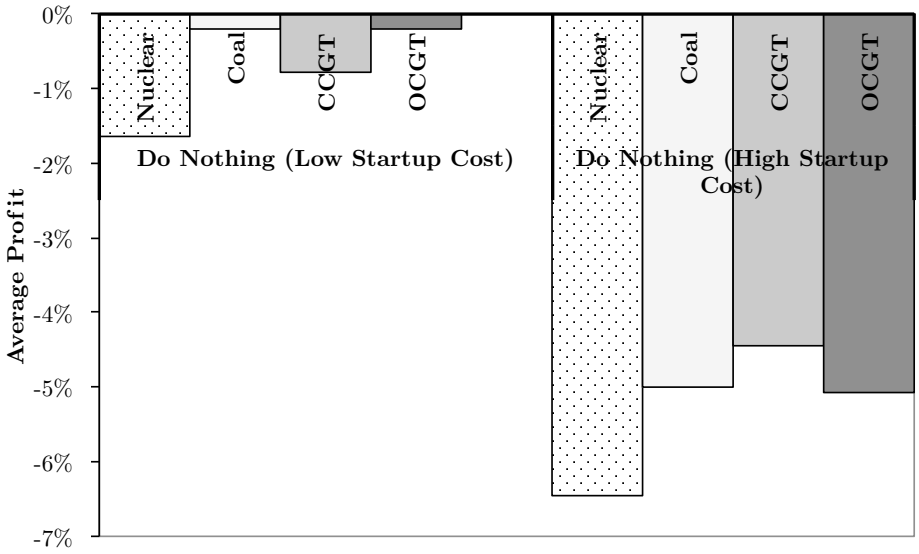


Figure 3-20: Sensitivity of technology-specific average profits to different values of startup cost, for the 'do nothing' case.

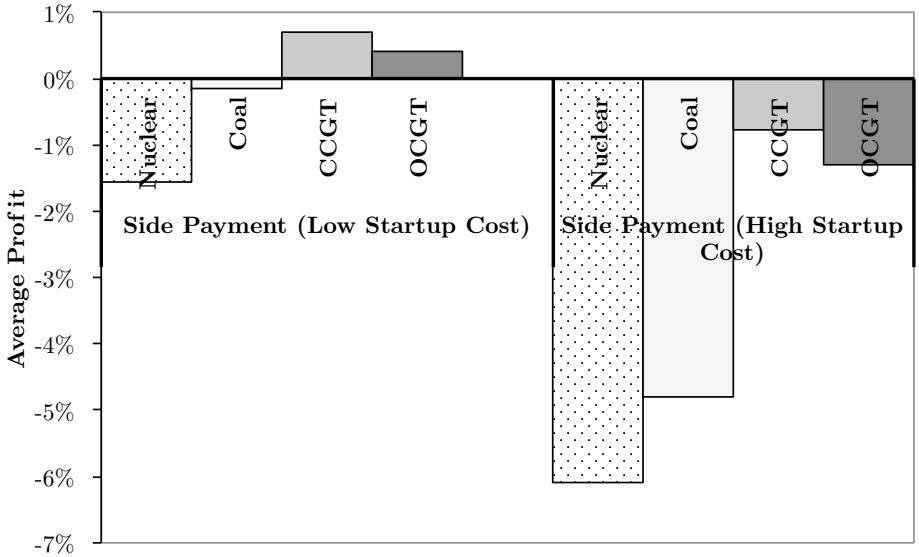


Figure 3-21: Sensitivity of technology-specific average profits to different values of startup cost, for the 'side payment case.

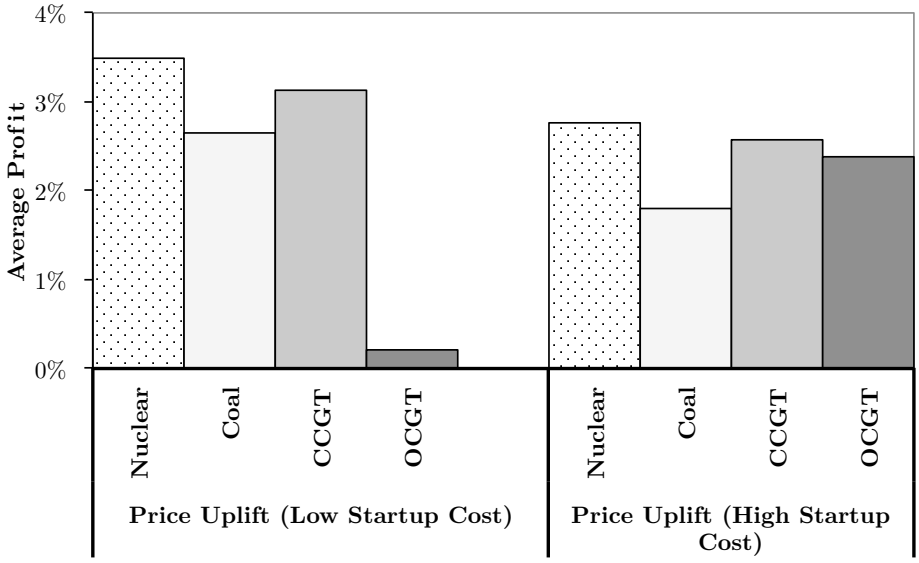


Figure 3-22: Sensitivity of technology-specific average profits to different values of startup cost, for the 'price uplift' case.

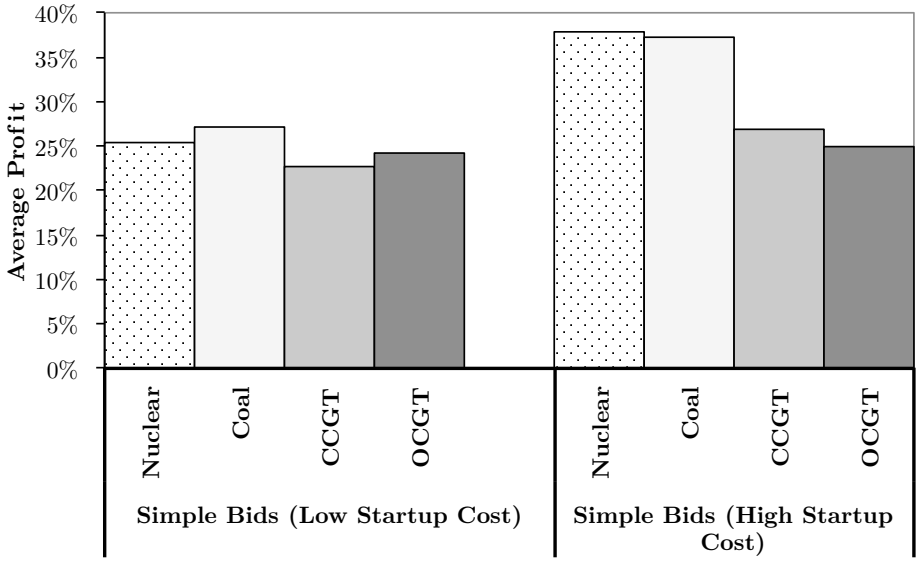


Figure 3-23: Sensitivity of technology-specific average profits to different values of startup cost, for the 'simple bids' case.

3.8.5 Impact of the Size of the Units

In addition to startup costs, another nonconvexity found in real systems is the lumpiness of the plants. Figure 3-24 shows the profits obtained in the original case together with those from case where the size of the units was reduced with respect to the size of the original plants (Appendix A). Plants in the pool of candidate plants were reduced as follows: nuclear plants were reduced by one half and one fifth, coal plants were reduced by one half and one fifth, CCGTs were reduced by one half and one fifth, and OCGTs were reduced by one fourth and one fifth.

Results in Table 3-24 show that as the size of plants is reduced, so is the the remuneration gap. This result, however, is not reflected on OCGTs because their startup cost is smaller compared to any other technology in the system, and these smaller OCGTs bear a more intense cycling regime than regular-sized ones. This increase in startup costs expressed in terms relative to a smaller unit size, make profit gaps larger.

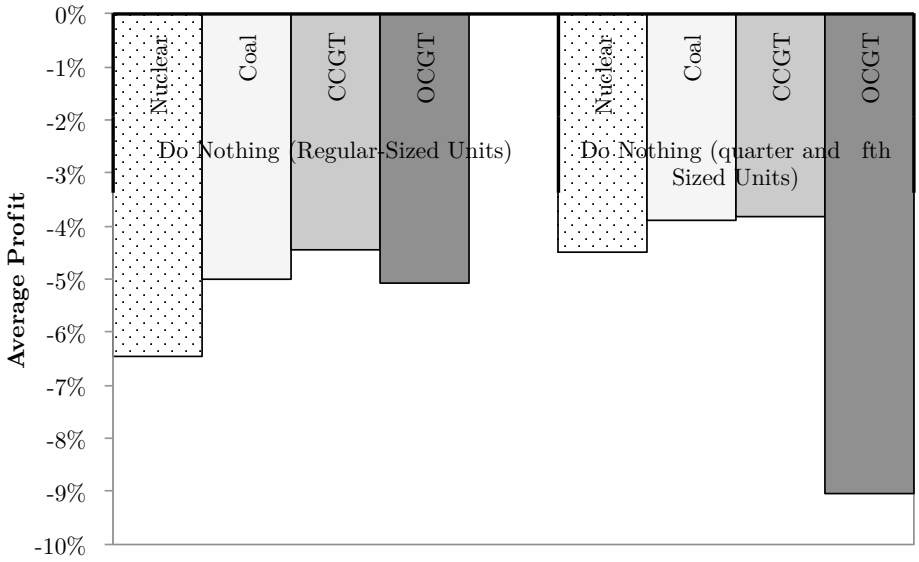


Figure 3-24: Sensitivity of technology-specific average profits to the size of units

3.9 Conclusion

The analyses in this chapter use IMRES, a capacity expansion model with embedded unit commitment constraints, to explore the impact of bidding rules in power systems with large shares of wind and solar PV, under two adaptation paradigms: a fully adapted generation fleet, optimized for the renewable scenario; and a fleet optimized for zero renewables operated with a certain non-zero renewable capacity.

Results show that bidding rules meant to ensure that plants recover nonconvex operation costs have significant implications in the overall cost recovery of the generation plants and the price paid by consumers. First, granting a side payment to power plants that do not have their operating costs covered by the wholesale price can bring them closer to full-cost recovery at very small cost to the consumers, but may fail to recover the total generation costs of the different technologies. Second, adding the optimal price uplift that allows all plants to recover operating costs increases the profits of all units, as startups increase the energy price through the uplift, which has a larger impact on the consumer and still may under-recover total generation costs. Third, allowing for the internalization of startup costs in simple bids grants plants revenues that exceed their total cost, creating opportunities for new entrants and imposing an unnecessary burden on the consumer.

In addition, this chapter explored the relative effects of bidding rules when there is an unanticipated deployment of renewables that impacts on the economic equilibrium of the system. Under this circumstance, an excess-capacity situation is created and generators incur losses in income due to a quick depression of wholesale prices and a reduction of their capacity factor, as a fraction of the demand is taken over by renewables. We found these losses to be significantly greater than those incurred from not recovering startup costs. This result underlines the criticality of having renewable energy plans that allow other players in the market to adapt without incurring losses provoked by potential regulatory shocks.

Chapter 4

Valuation of Risk at the Power Generation Plant

This chapter is an extended adaptation of the following paper:

- "A Dynamic Model for Risk Pricing in Generation Investments" (de Sisternes and Parsons, 2014).

4.1 Notation

This chapter uses a different notation than the one used in previous chapters of this thesis. The notation particular to this chapter is summarized in Table 4.1.

Table 4.1: Indices

$i \in \{0, 1, , N\}$, where i is the index the time periods
$j \in \{0, 1, , J(i)\}$, where j is the index for the uncertainty paths at a given period i
$t_i = T/Ni = i/n \in \{0, 1, , T\}$, where t_i is the time where period i is located

Table 4.2: Parameters

μ	expected market return over the course of one year

σ	expected volatility of the returns over the course of one year
α	expected demand growth over the course of one year
r	risk free rate
N	total number of periods in the simulation
T	total number of years in the simulation
$n = N/T$	number of periods per year
Δt	time duration of one period expressed in years
$m = \mu/n$	expected market return per period
$\nu = \sigma/\sqrt{n}$	standard deviation per period

Table 4.3: Variables

$\Delta R_{ij} \equiv \Delta R_{tj}$	market return at time period i and path j; $\Delta R_{ij} = \ln(S_{ij}/S_{i-1j})$
$R_{ij} \equiv R_{tj}$	cumulative market return at time period i and path j
$\varepsilon_{ij} \equiv \varepsilon_{tj}$	independently distributed random variable, $\varepsilon = \Phi^{-1}$
Φ	cumulative distribution function (CDF) of the standard normal distribution (with $\mu = 1$ and $\sigma = 0$)
$S_{ij} \equiv S_{tj}$	market price index at time period i and path j
$\pi_{ij} \equiv \pi_{tj}$	true probability of the state in time period i and path j
$\pi^*_{ij} \equiv \pi^*_{tj}$	risk neutral probability of the state in time period i and path j
$\phi_{ij} \equiv \phi_{tj}$	stochastic discount factor of the state in time period i and path j
$\bar{\lambda}$	technology-specific average risk premium
PV	present value of the asset
$ECF_i \equiv ECF_t$	expected cash flows of the asset, $ECF_\tau = \sum_{j=0}^{J} CF_{\tau j}\pi(R_{\tau j})$
$CEQ_i \equiv CEQ_t$	certainty equivalent value of the asset $CEQ_\tau = \sum_{j=0}^{J} CF_{\tau j}\pi^*(R_{\tau j})$
x_{ij}	electricity demand index in time period i and path j
$F(x_{ij})$	$\log x_{ij}$

4.2 Introduction

The previous chapter showed how conventional thermal technologies are exposed to the risk of unanticipated new renewable generation additions. Although the cost of wind and solar photovoltaics is steadily decreasing, these technologies are not yet cost competitive in many markets, making their deployment conditional on some sort of subsidy (feed-in tariffs, tax credits etc). These incentives that boost renewable deployment might be strongly correlated with the macro state of the economy, as programs that provide such economic incentives require large budgetary funds to be administered. If incentives to renewables are too aggressive and renewable deployment occurs unanticipatedly, the thermal generation capacity might not have enough time to adapt –think, for instance, of the lengthy operational life of a nuclear plant that might not be economical with a large renewable share–, affecting critically to the profitability of units.

The past number of years has seen significant interest in the role of risk in the valuation of electricity generating technologies. Companies are keenly aware that different projects contain different levels of risk, that project risk may vary throughout the project life-cycle, and that the contract terms negotiated for a specific project shift the risk dramatically. Policy makers, too, have looked to shape the risk of certain investments in order to steer the profile of new generation towards low carbon sources. Unfortunately, this heightened interest has not been matched with adequate tools for properly evaluating different risk profiles.

One approach leans on the now widespread availability of computing to generate large Monte Carlo distributions of payoffs to different assets or for the same asset financed with different contract. Usually the different distributions are compared on the basis of means and variances. For example, fixing the mean, a distribution with a higher variance is considered worse than a distribution with a lower variance.

One shortcoming of this approach is its failure to connect with the standard tools of modern valuation and asset pricing. This approach ignores the key insight from portfolio theory that expected return is not a function of total variance, but rather of the component of variance that is correlated to macroeconomic variables. It also ignores the key insight from

derivative pricing that variance in the final payoff is a poor tool for ranking risk. The non-linearity of many payoffs makes the problem more difficult than is acknowledged in a simple mean-variance framework. This disconnect undermines the reliability of many conclusions drawn from these Monte Carlo simulations, and it undermines the confidence we might have in the specific values calculated using the simulations.

This chapter shows how to incorporate standard risk pricing principles into the popular Monte Carlo simulation analysis. The method presented is basically a version of what is commonly known as the 'real options' approach to asset valuation (Dixit and Pindyck, 1994)(Cox et al., 1979). Another label for it is 'contingent claims valuation'. Unfortunately, for various reasons these tools are not yet widely employed by financial and policy analysts working in the power industry. One reason may be the difficult structure of those approaches and the lack of any obvious connection to the Monte Carlo simulations as most often employed by financial analysts in the power industry. This chapter attempts to take a different approach designed to remedy that problem. The exposition is structured using the standard framework for a typical Monte Carlo cash flow simulation so that the implementation can be readily generalized. The framework proposed helps to make clear what the necessity of an asset pricing approach places on the financial analyst: it is essential that the analyst assess the relationship between the risk in the asset's cash flows and the macroeconomic risk that financial investors are concerned with. Without that, it is impossible to properly think through how an assets risk is priced in the market, and equally impossible to assess how different contracts or other changes impacting the asset's risk will be valued. The framework provided is flexible and can accommodate many different structures for the relationship between the risk in the asset's cash flows and the macroeconomic risks. This chapter is intended to help remedy that.

There have been attempts in the literature to account for the non-linear exposures of cash-flows in certain assets, and to derive discount rates that reflect those exposures. For instance, outside the realm of the purely financial industry, Myers and Howe account for the non-linearity of R&D costs in the pharmaceutical industry and how that reflects on

the cost of capital across different development stages of a drug (Myers and Howe, 1997). However, this work does not extract the component of the variance that is correlated with macroeconomic variables in order to appropriately price risk. Conversely, Ehrenmann and Smeers establish a relationship between the market return and the realizations of a group of stochastic parameters that affect cash flows (Ehrenmann and Smeers, 2010). They derive a discount factor for each scenario via the CAPM, which are later used in a two-stage electricity generation capacity expansion equilibrium problem. Nevertheless, this work does not account for the dynamic evolution of the underlying risk factor (for instance, as a random walk or some other stochastic process) and its effect on cash flows.

4.3 The General Algorithm

This section reviews a couple of principles of valuation and asset pricing that need to be introduced in the analysis of risk in generation assets. In doing so, this section defines a general algorithm that can be employed across a wide range of problems. An illustration of how the steps of this algorithm are implemented is also provided. The illustration reveals how the algorithm can encompass complicated risk structures, while simultaneously demonstrating how the algorithm ties back to the most basic discounted cash flow exercises.

Most modern asset pricing models value a projects risky cash flows with reference to the larger economic context of which the individual project is just a small part. This happens in two steps. First, the larger context defines a standard unit of risk and the market price for that standard unit. Second, the project is analyzed to determine how much of the standard unit of risk is contained in the project, and then the price of risk for that project is the product of the amount of risk measured in those standard units and the market price of risk per standard unit.

For example, the Capital Asset Pricing Model (CAPM) looks to (i) the volatility of a well diversified market portfolio of stocks and other securities and assets, and (ii) the expected return premium earned on this market portfolio. The market portfolio is the standard unit of risk, with a Beta of 1, and the premium is the price of risk paid per unit of Beta risk. Then,

in evaluating a project, we determine the projects Beta –i.e., how much of the standard unit of risk are contained in the project– and calculate a risk premium for the project equal to the projects Beta multiplied by the market risk premium.

Methodologies like the CAPM work well for projects with risk that is simply proportional to the market risk. When a projects risk is non-linearly related to market risk, a similar but more complex algorithm is required. These more complex algorithms employ what is known as a stochastic discount factor, which is what this chapter presents. But these more complex algorithms still operate in the two steps described above. First, the algorithm identifies the market price of a standard unit of risk, and, second, the algorithm determines how much of this standard unit of risk is contained in the project.

In the algorithm presented, this first step is called the Macro Perspective, and the second step is called the Micro Perspective. The analysis looks first away from the project and to the larger economic context to identify how the pertinent risks are priced in the market. Then, the analysis looks back to the project to identify how much of the pertinent risks are contains in the project. Once these two steps are done, we can execute the valuation of the project and conduct any number of analyses (Figure 4-1).

While a final analysis can be presented as being executed linearly, starting with the macro analysis, moving on to the micro analysis, and then executing the valuation and any further analysis, the actual work leading towards the final analysis will involve an iterative process operating sometimes in one direction and sometimes in another. For example, depending upon the project being analyzed, we may choose to focus on a different choice for the standard unit of market risk.

It is appropriate now to formalize this discussion and recapitulate the points above inside that formal framework. Let's suppose we have a project that operates over a number of years, $t = 1, T$. The traditional discounted cash flow formula is given by the expression:

$$PV = \sum_{t=0}^{T} \frac{\mathbb{E}[CF_t]}{(1 + R_a)^t} \tag{4.1}$$

For our purposes, it will be important to rewrite the same equation using the correspond-

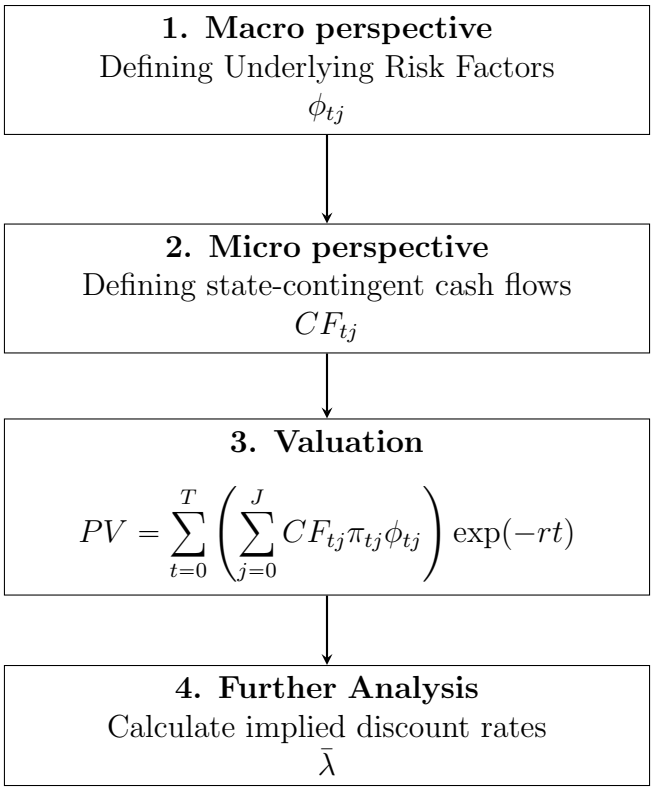


Figure 4-1: Steps of the General Algorithm.

ing continuous time discount rate (see Appendix B):

$$PV = \sum_{t=0}^{T} \mathbb{E}[CF_t] \exp\left(-r_a t\right) \tag{4.2}$$

–where $r_a = \ln\left(1 + R_a\right)$, and then to decompose the discount rate into the risk-free rate, r, and a risk premium, λ_a:

$$PV = \sum_{t=0}^{T} \mathbb{E}[CF_t] \exp\left(-\lambda_a t\right) \exp\left(-rt\right) \tag{4.3}$$

In the capital asset pricing model (CAPM) (Copeland et al., 2005), the appropriate risk premium, λ_a, is the product of the correlation between the historical return of the security and the market portfolio (β), and the market risk premium $r_p = (r_m - r)$, where r_m is the return of the market portfolio and r is the risk-free rate of return. As a result, the required

rate of return on asset "a" is $r_a = r + \lambda_a = r + \beta_a\ r_p$, so that we have:

$$PV = \sum_{t=0}^{T} \mathbb{E}[CF_t] \exp\left(-\beta_a r_p t\right) \exp\left(-rt\right) \tag{4.4}$$

In our terminology, identifying the market risk premium, r_p, is the macro task and determining the asset Beta, β_a, is the micro task. Once these two tasks are complete, the valuation is executed according to the equation[1].

This standard discounting formula is simple in that it employs the expected cash flow, and a single risk adjusted discount rate. However, it is often forgotten or overlooked that this formula is a special case of a more general discounting method that employs the full space of possible outcome states determining the cash flows, and applying different discount factors to the cash flows in different states – see, for example, (Cochrane, 2005). The set of different discount factors is known as the stochastic discount factor. The validity of reducing the more general method to the standard discounting formula rests upon two assumptions that are rarely met. First, it assumes that the risky cash flows being discounted are perfectly proportional to the underlying risk factor so that the risk can be completely summarized by a scalar such as β. This assumption assumes away the existence of price caps, price floors and other nonlinearities which would make it impossible to rank assets by risk along a single dimension. Second, the risk in the cash flows grows linearly through time. This assumption will be violated if uncertainty is unevenly distributed over different horizons, as, for instance, in projects where much of the uncertainty is resolved in early stages. In the context of electricity generation, these two assumptions would entail ignoring, among others, the relationship between the general state of the economy and electricity demand growth, and the of regulatory changes to the stream of cash flows.

The stochastic discount factor methodology presented here extends the standard discounting formula to handle cases in which one of these two assumptions no longer holds. It

[1]The CAPM is just one possible model for determining the market price of risk. Other models include the Fama-French 3-factor model or the Arbitrage Pricing Model, to name a couple. And, derivative pricing models exist which start from the market price of risk embedded in an underlying security, and then move on to evaluate derivative securities with payoffs linked to the payoff on the underlying security.

does this by making explicit the cash flows exposure to the underlying risk factor and by discounting future cash flows at rates that are contingent on the realizations of the uncertain factor. Assuming that exposure is well-defined, the method involves determining the value of the risk factor in all possible future states and, based on that value, forecasting the cash flows that the asset will generate in each state. Accordingly, risk is priced through explicitly taking into account the correlation between the cash flows and the underlying risk factor.

Suppose that in any year t, the underlying priced risk factor can take on a discrete range of values, $j = 1, J$. Assume that each possible state has a probability associated (π_{tj}), equal to the probability of the realization of the risk factor, such that $\sum_{j=0}^{J} \pi_{tj} = 1, \quad \forall t$. In addition, for each possible state of the underlying priced risk factor, we know the project cash flows contingent on that risk factor, CF_{tj}. Then the present value formula is given by Eq. .

$$PV = \sum_{t=0}^{T} \left(\sum_{j=0}^{J} CF_{tj} \pi_{tj} \phi_{tj} \right) \exp(-rt) \tag{4.5}$$

–where ϕ_{tj} is a state contingent discount factor. Here is where the difference lies between risk-adjusted discounting and state contingent discounting: while in the former method cash flows are discounted using a fixed factor ($e^{-\lambda}$) that is common to all time periods and uncertainty states, in the latter there is a different discount factor ϕ_{tj} for each state. In both cases, the discount factors are taken from the market and are independent of the cash flows.

It may be helpful to illustrate the general methodology with a simple numerical example. Then, the remainder of the chapter produces a detailed and more realistic implementation with a clear tie to alternative electricity generating technologies. Table C-1 shows a standard discounted cash flow valuation of PowerCo, a generator with 5 years of operations. The risk-free discount rate is constant through time at 5%, the risk-premium is constant at 3%, and the company has a Beta of 2.0. So PowerCo's risk premium is 6%, and, adding the risk-free rate the complete risk-adjusted discount rate is 11%. Table C-1 repeats the calculation a second time, but breaking the discounting up into two steps according to the certainty-equivalent method. The first step does the discounting for risk using just PowerCo's risk premium of 6%, and the second step then does the discounting for time using the risk-free rate of

5%. In the following steps we show how to repeat this calculation using the more detailed stochastic discount factor methodology. This section concentrates at first on replicating this calculation, and then it shows how the stochastic discount factor methodology is more robust to variations on PowerCo's risk structure.

4.3.1 The Macro Perspective: Defining the Underlying Priced Risk Factors

Table C-2 unpacks the probability distribution of the underlying risk factor for our example. The most widely used underlying risk factor in corporate finance is the return on the market, i.e., the return on a broadly diversified index of financial assets, usually stocks. That is the risk factor used in this example. Let's assume that the return on the market index is normally distributed with a mean annual return of 6.4%, and an annual volatility of 18%. This example approximates the probability distribution of the cumulative return at each horizon using a binomial probability distribution defined on six possible outcomes or "states", indexed 0-5. At the one year horizon, the table shows a market return that ranges from a high of 47% to a low of -34%. Assuming that the market index at time $t = 0$ is set to 1, we can calculate the corresponding market index values. At the one year horizon, the table shows a market index value that ranges from a high of 1.594 to a low of 0.713. The binomial probability distribution across these six outcomes appears next.

The table shows a range of cumulative market returns and market index values at five time horizons, $t = 1, 2, , 5$. The probability distribution at each horizon is from the perspective of $t = 0$, and each distribution has been approximated separately. The reader is cautioned against thinking there is any tie between the set of states shown for the realized market return in year 1 and the set of states shown for the realized market return in year 2. The reader should think of the six states on each distribution separately. These are not six scenarios tracked through time, but six possible outcomes at each point in time, and the actual market index might arrive at any later outcome by many different paths through time. Since the uncertainty about the market return is growing through time, the range

of the market index from the highest point of the distribution to the lowest point of the distribution is also growing through time. Since the table shows returns on an annualized basis, this is less obvious when looking at the range of annualized returns.

This is all relatively standard fare. The new elements are the state-contingent risk premia shown, and the corresponding state-contingent discount factors. Instead of a single risk premium of $r_p = 8\%$, we have a set of risk premia. In each year, each state has its own risk premium. Some of these risk premia are very large–more than 30%! Others are even negative, which implies that someone would pay to accept that risk. For the purposes of illustrating the mechanics of the general algorithm, the state contingent discount factors are provided without showing how to derive them. The following sections of the chapter develop a fuller model of this methodology applied to electricity generating investments, showing how to derive the state contingent discount factors.

4.3.2 The Micro Perspective: Defining the State-Contingent Cash Flows for an Asset

Table C-3 unpacks how PowerCo's revenues are exposed to the market risk factor. Production is assumed to be independent of the market fluctuations, so that all variations in the revenue are due to changes in the unit price. For each state of the market index, Table C-3 shows a corresponding realization of PowerCo's unit price and revenue. The probabilities shown are the probabilities of the states, i.e. those shown in Table C-2. Let the subscript j denote a specific outcome or state, with p_{tj} denoting the unit price in state j at time t, and π_{tj} denoting the probability of state j at time t. Then, the expected unit price is the sum product of the state contingent unit price and the probabilities:

$$\mathbb{E}[p_t] = \sum_{j=0}^{J} p_{tj}\pi_{tj} \tag{4.6}$$

It matches the expected unit price used in Table C-1 to calculate the standard discounted cash flow value. Denoting with CF_{tj} the revenue in state j at time t, the expected revenue

is the sum product of the state contingent revenue and the probabilities:

$$\mathbb{E}[CF_t] = \sum_{j=0}^{J} CF_{tj}\pi_{tj} \tag{4.7}$$

It also matches the expected revenue used in Table C-1 to calculate the standard discounted cash flow value.

4.3.3 Executing a Valuation

Table C-4 shows how we use the new information on stochastic discount factors to calculate the present value of PowerCo's revenue over the next five years. We calculate the full range of possible revenues, multiply it times the respective probabilities, multiply that times the respective discount factors, and then take the summation across states. This gives us the certainty-equivalent revenue:

$$CEQ_t = \sum_{j=0}^{J} CF_{tj}\pi_{tj}\phi_{tj} \tag{4.8}$$

Once we have the certainty-equivalent revenue, we discount it for time using the risk-free discount rate:

$$PV_t = CEQ_t \exp -rt \tag{4.9}$$

Although this calculation involves more detail on the full range of possible outcomes for the underlying risk factor and PowerCo's revenues, it yields the same final present value.

The simpler discounted cash flow calculation in Table C-1 uses the expected cash flow. The calculation shown in Table C-4 replaces that with an explicit specification of the scenarios that went into the expectation, and with a correspondingly explicit set of discount factors:

$$PV = \sum_{t=0}^{T} \left(\sum_{j=0}^{J} CF_{tj}\pi_{tj}\phi_{tj} \right) \exp\left(-rt\right) \tag{4.10}$$

4.3.4 Analysis

The value of using state-contingent discount factors is that we can now very easily value any variation on the original payoff structure that is also contingent on the underlying risk factor. Suppose, for example, that PowerCo is owned as a partnership, and that revenues are divided up according to the following sharing rule: partner #1 receives the first \$0.37 per unit sold, while partner #2 receives any revenue over and above that. Tables C-5 and C-6 show how these two partnership positions are valued using the state-contingent discount factors. Partnership #1 is worth \$10.201 billion, while Partnership #2 is worth \$32.254 billion. The value of the two positions sum to the total value of the project, of course.

The two partnership stakes have very different risk exposures. Partner #1 has a nearly riskless position. As one can see in Table C-5, only in years 3, 4 and 5, and then only in the event of the lowest possible price, does partner #1 receive something less than \$0.37 per unit. Otherwise, the price received by partner #1 is constant regardless of the realization of the risky unit price in the market. Indeed, we could have almost exactly valued partner #1's position using the expected cash flow and a riskless discount rate:

$$PV(\#1) \approxeq \sum_{t=0}^{5} (\mathbb{E}_0[CF_t]) \exp(-rt) = \$10.221. \tag{4.11}$$

This is only slightly above the correct valuation made using the state contingent valuation methodology.

However, we could not very easily have applied the DCF technique to value partner #2's share. Why not? The problem that arises is how to choose the right discount rate. All of the project's risk has been loaded onto partner #2's share of the revenue. Look at the range of revenue received by partner #2 as shown in Table C-6. The range is almost as wide as the range for the total firm, even though partner #2's stake is less than the value of the whole firm. On the upside, partner #2's realized rate of return will be greater than that of the firm as a whole, and on the downside it will be lower. Clearly the old risk premium of $\lambda_a = 6\%$ that was used for the project as a whole is not the right number to use for partner

#2's share. If we tried to calculate the partnership value using the 6% discount rate, we would significantly exaggerate the its value:

$$PV(\#2) \neq \sum_{t=0}^{5} (\mathbb{E}_0[CF\#2_t]) \exp(-6\%t) \exp(-rt) = \$37.581. \tag{4.12}$$

This overvalues partner #2s stake by 17%.

The right risk premium must be much higher, to reflect the loading of extra risk. But how much higher? The simplest way to answer that question is to back out the discount rate that gives the same answer as the calculation made with the state-contingent discount factor.

The methodology of applying state-contingent discount factors is very flexible. The hard part is specifying the array of states and the state-contingent discount factors, as shown in Table C-2, and tying these to the cash flows on the asset we want to value, as shown in Table C-3. Once this has been done, we are free to evaluate all variety of generation technologies with different risk exposures.

The general algorithm is now applied to the problem of pricing risk in generation assets. To do this the next section first defines the model of the underlying priced risk factor. It then overlays on top of that a model of electricity price determination and its correlation to the underlying priced risk factor. Finally, it defines specific electricity price related assets which we want to value, and describes how their cash flows are determined.

4.4 The Underlying Priced Risk Factor: the Stock Market Index

4.4.1 The Probability Distribution of Market Returns

The most widely used model of risk is the Capital Asset Pricing Model (CAPM) in which the return on a diversified portfolio of stocks is the single priced risk factor. Consistent with that model, let's assume that the cumulative return on the market portfolio, R, evolves as a

random walk or arithmetic Brownian motion. The stochastic process for R is described by two parameters: the drift, μ, and the volatility, σ:

$$dR = \mu dt + \sigma dW \tag{4.13}$$

– where dW is the increment to a Brownian motion.

Let's denote by $\pi(R_{\tau j})$ the probability distribution of the cumulative market return at time horizon τ and uncertainty path j. Given the Brownian evolution of the cumulative market return, this probability is given by a normal distribution function with mean $\mu\tau$ and variance $\sigma^2\tau$. If we denote by $f_1(\cdot)$ the p.d.f. of the normal distribution $\mathcal{N}(\mu\tau, \sigma\sqrt{\tau})$, and by J the number of states sampled in that time period, the true probability of a given state relative to the other states sampled is given by the following expression:

$$\pi(R_{\tau j}) = \frac{f_1(R_{\tau j})}{\sum_{j=0}^{J} f_1(R_{\tau j})} \tag{4.14}$$

Figure 4-2 shows the mean of this distribution through time, together with an upper and lower confidence bound. The figure also shows the full distribution at one slice of time. Overlayed on top of the figure is one possible sample path for the cumulative market return.

4.4.2 The Stochastic Discount Factor

Let's denote by $\phi(R_{\tau j})$ the stochastic discount factor to be applied to cash flows earned at time horizon τ and uncertainty path j, in the state defined by cumulative market return $R_{\tau j}$. Already, because of the strong assumption on how the market returns evolve, we have implicitly determined the stochastic discount factors that must be used. At each horizon, τ, the function created by multiplying the probability of the cumulative market return, with the stochastic discount factor for that return, $\pi^*(R_{\tau j}) = \pi(R_{\tau j})\phi(R_{\tau j})$, has the features of a probability distribution –if the distribution were defined on a discrete state space, then those properties would be (i) the probability of every event lies between 0 and 1, and (ii) the sum of the probabilities equals 1. The resulting distribution is often called the 'risk-neutral'

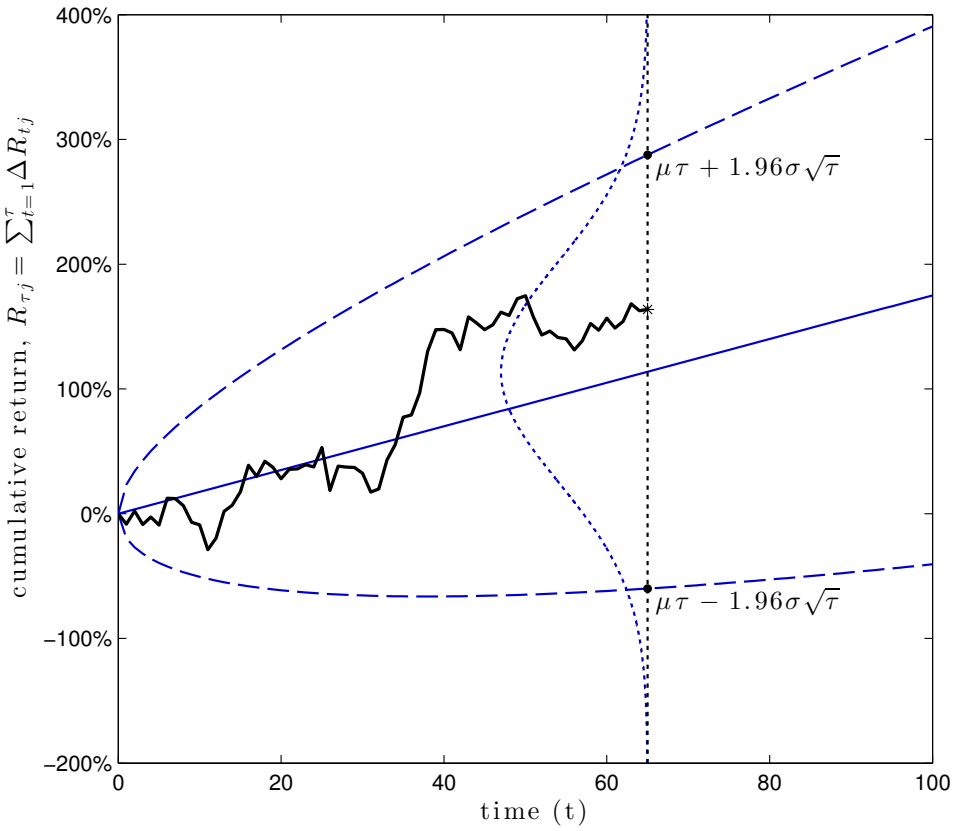


Figure 4-2: Graphical representation of the true probability distribution of market returns.

distribution[2]. Given the assumption about the evolution of the market return, the resulting distribution must also be normally distributed, with mean $(\mu - \lambda)\tau = r\tau$ and variance $\sigma^2\tau$, where r is the risk-free rate, μ is the total expected market return and $\lambda = \mu - r$ is the market risk premium. If we denote by $f_2(\cdot)$ the p.d.f. of the normal distribution $\mathcal{N}(r\tau, \sigma\sqrt{\tau})$, and by J the number of states sampled in that time period, the risk-neutral probability of a given state relative to the other states sampled is given by the following expression:

$$\pi^*(R_{\tau j}) = \frac{f_2(R_{\tau j})}{\sum_{j=0}^{J} f_2(R_{\tau j})} \tag{4.15}$$

Figure 4-3 shows the mean and confidence bounds of this risk neutral distribution overlayed on the mean and confidence bounds of the true distribution. The figure also shown

[2]see for instance (Duffie, 2001) for a rigorous treatment of risk neutral valuation

the full risk-neutral distribution at one slice of time, overlayed on the true distribution at that slice of time.

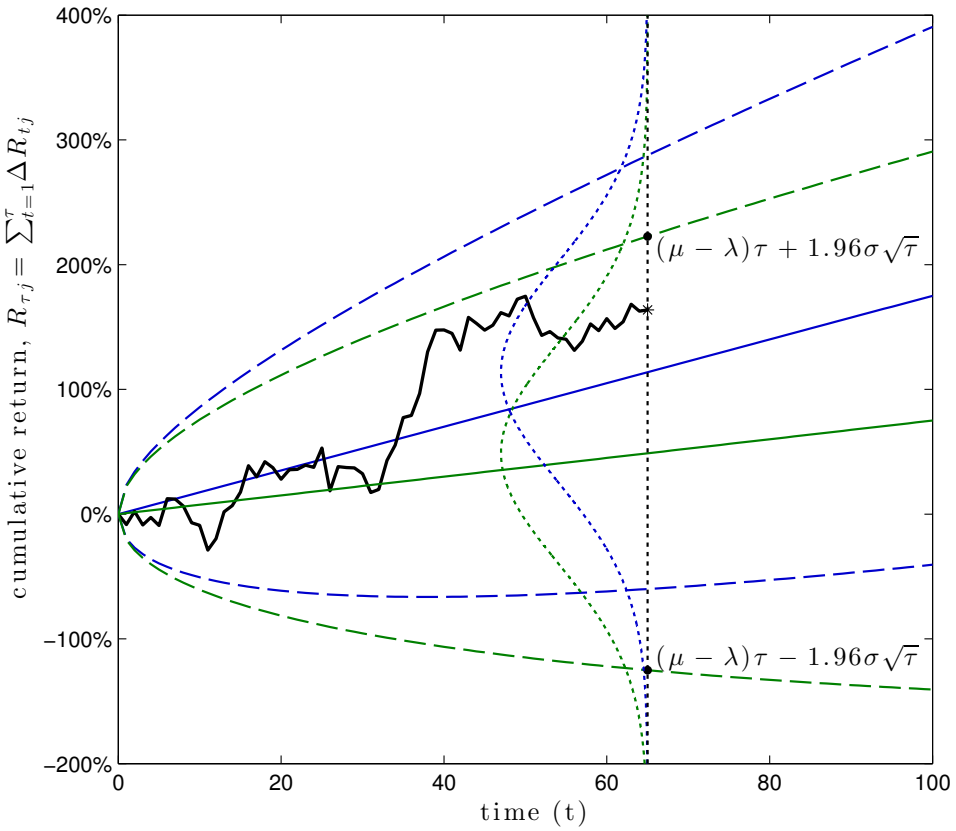


Figure 4-3: Graphical representation of the risk-neutral probability distribution of market returns.

Given the true probability distribution, and the risk-neutral distribution formed as the product of the true distribution and the stochastic discount factor, we can back out the stochastic discount factor if necessary. However, for all of our valuation work, what we need is the product of the true probability and the stochastic discount factor. That is, what we need is the risk-neutral probability distribution. So it is not necessary to recover just the stochastic discount factor. Nevertheless, one could produce it if desired.

The state-contingent discount factor can thus be obtained by dividing the risk neutral

probability and the true probability:

$$\phi(R_{\tau j}) = \frac{\pi^*(R_{\tau j})}{\pi(R_{\tau j})} \tag{4.16}$$

Intuitively, it can become apparent how the risk-neutral distribution accounts for both the probability of the states and their risk: if one compares the point marked with an asterisk in Figure 4-2 and the same point in Figure 4-3, the probability of that same state is larger in the true probability distribution than in the risk-neutral distribution. Therefore, for the case shown in the figure, $\phi(R_{\tau j}) < 1$, meaning that the risk premium associated with this state is positive, and that investors will value cash flows in this state less than if those same cash flows were received at present.

4.5 A Model of Electricity Prices

In order to apply the stochastic discount factor, we need to a model that ties the cash flows of any asset being valued back to the underlying risk factor. Therefore, at least one link between both needs to be established in order to build that connection. In this implementation let's assume that demand growth is directly correlated with the market index and that demand has an effect on the electricity price. Other more complex relationships could have also been established between the risk factor and other variables including regulatory shocks or fuel prices.

Let x be an index parameter capturing the electricity demand at a certain state, so that the evolution of x represents the evolution of electricity demand through time (it represents changes in demand, but its value is not equal to total demand). x can be thought of as a simplified representation of demand, that assumes a certain profile with some seasonality, where the seasons evolve in the same proportions as the demand index. Let's also assume that the stochastic process describing the growth of electricity demand is governed by a *balanced-growth* model, where demand growth is perfectly correlated with the underlying

priced risk factor, albeit with a different drift, α:

$$\frac{dx}{x} = \alpha dt + \sigma dW \tag{4.17}$$

–where σ is the volatility of the cumulative market return and dW is the increment to a Brownian motion describing the market return.

Let p be an electricity price index, abstracting from fluctuations associated with daily, weekly and seasonal cycles. This price index can be thought of as the profits of an industry whose growth is evolving through time. Let q be an index of total installed generation capacity, reflecting a standard portfolio of capacity types including baseload, peakers and units providing a variety of whatever capacity and ancillary services may be required. The price is linked to the level of demand and the level of installed capacity through an inverse demand function with a constant elasticity $-1/\gamma$. Formally:

$$p = D(q, x) = xq^{-\frac{1}{\gamma}} \qquad [\$/\text{kW-period}] \tag{4.18}$$

As demand outgrows installed capacity, price rises. Let's assume that at some level of price, p^*, new capacity is installed instantaneously, so that the trigger price, p^*, is effectively a price cap. Consequently, the electricity price becomes a 'regulated Brownian motion', inheriting volatility from demand and the underlying priced risk factor, but not reflecting that risk perfectly.

The trigger price equilibrium is inspired by Leahy (Leahy, 1993). If we consider that a representative portfolio of capacity has both fixed and variable costs, the total costs of the new entrant (I) is the sum of the present value of its cost of capital, K, plus the present value of its variable costs, C/r, where C is considered a constant annuity equal to the variable cost of the plant, c, times 8,760 hours, times the average utilization of the plant, u:

$$I = K + C/r \qquad [\text{k\$/MW}] \tag{4.19}$$

$$C = c \cdot 8,760 \cdot u \qquad [\$/\text{kW-yr}] \tag{4.20}$$

We now calculate the 'flow equivalent' (per unit of time) cost of investment, δI, which is the level the initial price should have if its subsequent value is to recover the cost of investment (Dixit and Pindyck, 1994), where δ is equal to the risk free rate per period (r) minus the demand growth rate per period (α). Capacity is assumed to have no depreciation. At the trigger price level ($p^* > \delta I$), the value of the option to invest, $F(p^*)$, is equal to the value of exercising the option, $V(p^*) - I$. This is the 'value matching' condition found in (Dixit and Pindyck, 1994). This calculation considers also no depreciation of the asset, assuming that once the asset is installed it keeps producing the output flow forever. Here the trigger price p^* represents the net revenue that a unit will receive for generating one megawatt throughout the period considered (one year, or a fraction of the year), and it is the price at which the total generating capacity of the system will be increased, and making prices to subsequently fall. Following Dixit and Pindyck, the optimal price trigger entry for a new entrant is:

$$p^* = \frac{\beta}{\beta - 1}\delta I \qquad [\$/\text{kW-period}] \tag{4.21}$$

Here, the factor $\frac{\beta}{\beta-1} > 1$ is the 'option value multiple' and reflects the price should be high enough to account also for the value of waiting to invest and the 'option value multiple' (uncertainty and irreversibility of the investment), with β being the positive root of the polynomial:

$$\beta = \frac{1}{2} - \frac{\alpha}{\sigma^2} + \sqrt{\left(\frac{\alpha}{\sigma^2} - \frac{1}{2}\right)^2 + \frac{2r}{\sigma^2}} \tag{4.22}$$

This approach does not account for the fact that operation can be temporarily and costlessly suspended when $p < C$, and costlessly resumed when $p > C$.[3] The trigger price (i.e., the net revenue obtained during one period) can be easily translated into hourly price units by simply dividing the revenue obtained during that period by the energy produced by the plant, by means of its utilization factor u:

$$\rho^* = \frac{p^*}{8,760 \cdot u} n \qquad [\$/\text{MWh}] \tag{4.23}$$

[3]This feature is accounted for on page 186 of (Dixit and Pindyck, 1994)

4.6 Valuing Assets With Cash Flows Tied to the Electricity Price

Having defined the electricity price index and how it reflects the risk of the underlying priced risk factor, and armed with the stochastic discount factors associated with this underlying priced risk, we are now ready to value any asset with cash flows that are tied to the electricity price. Let's focus on three different assets.

The first is a pure financial asset, an electricity price swap which pays a cash flow equal to the electricity price index on 1 MW. The risk of this financial asset represents the risk intrinsic to the electricity price.

The second asset is a call option on the electricity price with a price strike equal to c_b, where c_b is the marginal cost of an existing unit of generation capacity with a low marginal operating cost (a baseload unit). Therefore, the risk of this option is equivalent to the risk of an existing baseload plant. This power plant has a simple operating strategy in which it is on and generating whenever the hourly electricity price is above marginal cost, and off whenever the hourly electricity price is below. Care must be taken here in translating between the hourly price earned and the industry index price which is an average across the particular calendar patterns of the hourly price. Given the baseload units marginal cost, operating strategy and the calendar pattern, its average number of hours of operation within a seasonal cycle will vary with the level of the index price, as will the average price earned in that cycle. We oversimplify the problem and assume that the cash flows earned by the option equal to $\max\{\rho - c_b, 0\}\frac{8{,}760}{n}$.

The third asset is another call option on the electricity price with an strike price equal to c_p, where c_p is the marginal cost of an existing unit of generation capacity with a high marginal operating cost. We call this the peaker unit. This unit, too, has the simple operating strategy in which it is on and generating whenever the hourly electricity price is above marginal cost, and off whenever the hourly electricity price is below. The same proviso applies about translating between the hourly price earned. We oversimplify the problem and

assume that cash flows earned by the option equal to $\max\{\rho - c_p, 0\}\frac{8{,}760}{n}$.

Each of these assets can be valued directly with explicit formulas derived using the continuous time mathematics. However, the objective of this chapter is to show how to use Monte Carlo simulation to perform a valuation, so we turn now to an implementation of the valuation model using a Monte Carlo simulation.

4.7 Implementation Via Monte Carlo Simulation

The risk neutral valuation method is applied to a two-technology power system (baseload and peaker), with 50GW of initial installed capacity in the system, and with demand growth perfectly correlated with the market return index. The values of the parameters used in the model are given in Table 4.4.

Table 4.4: Parameters used in the implementation

Parameter	Value	Description
$\mu - r$	0.04	Annual market risk premium
σ	0.22	Annual market volatility
r	0.03	Annual risk free rate
α	0.02	Annual drift rate for demand
β	1.204	Investment opportunity parameter
I	\$7708/kW	Present value of costs of new entrant
p^*	\$113.936/kW-period	Leahy price trigger
ρ^*	\$86.709/MWh	Price trigger in hourly price units
γ	1.2	Inverse demand elasticity
q_{0j}	50 GW	Initially installed generating capacity
x_{0j}	$20q_{0j}$	Initial demand shock
c_b	$0.1\rho^*$	Baseload maginal cost
c_p	$0.9\rho^*$	Peaker maginal cost
P_b	1 GW	Baseload unit capacity
P_p	0.2 GW	Peaker unit capacity

In our Monte Carlo simulation, time periods are indexed by $i \in \{0, 1, , N\}$, where $N = nT$

is the total number of periods in which the time horizon of T years is divided. Each time periods has length Δt years, with $\Delta t = 1/n$. Sample paths are indexed by $j \in \{0, 1, , J\}$. We run a 50-year simulation, using 4 periods per year. We calculate 100-sample paths.

4.7.1 Evolution of the Underlying Priced Risk Factor: the Stock Market Index

The discrete time representation of the underlying risk factor (the market return, ΔR_{ij}) is: $\Delta R_{ij} = m + \nu\varepsilon_{ij} = \mu\Delta t + \sigma\sqrt{\Delta t}\varepsilon_{ij}$, where $m = \mu/n$ is the expected market return per period; $\nu = \sigma/\sqrt{n}$ is the standard deviation per period; and the step ε_{ij} is a independently distributed random variable with a mean of zero and a standard deviation of one. In addition, let's assume that these steps are normally distributed ($\varepsilon_{ij} = \Phi^{-1}$) and sample possible uncertainty paths for each time period using the Monte Carlo method (Rubinstein and Kroese, 2008):

- For each time period i, we draw a Latin Hypercube (McKay et al., 1979) sample $\boldsymbol{h}_i \in (0, 1)^J$, where $h_{ij} \sim U(0, 1)$. Latin Hypercube Sampling is a variance reduction technique that is based on the idea of stratified sampling. In this approach, the interval is divided into J equally probable subintervals, and a uniform sample $\boldsymbol{h}_i = \{h_{i0}, ..., h_{iJ}\}$ is generated from within each subinterval.

- With each h_{ij} we now apply the inverse-transform method to draw a sample from a normal distribution of mean zero and standard deviation one. In this way, we can generate the normally distributed variable ε_{ij} (the error term of the random walk) by calculating the inverse of the standard normal cumulative distribution (cdf) function: $\varepsilon_{ij} = \Phi^{-1}(h_{ij})$.

The true probability of one state at a given time period, π_{ij}, is the probability of the cumulative return at state i, $R_{ij} = \sum_{k=1}^{i} \Delta R_{kj}$, through some certain path j. The distribution of all possible outcomes in one period is normally distributed with mean the expected cumulative return (as the mean value of the cumulative return grows with time), and a standard deviation equal to the standard deviation of market returns times the square root of time. In

other words, the true probability is the probability of a certain cumulative return outcome, given the distribution of all possible outcomes at that period. This distribution changes from one period to the next as a consequence of market returns evolving as a random walk.

Note also that the probability obtained is expressed relative to all the possible realizations in that period. Instead, we are interested in expressing this probability relative to the realizations sampled in that period. Therefore, as the distribution is being approximated by the samples drawn, we need to normalize the probabilities obtained by the sum of the probabilities of the other states sampled in the period.

The risk-neutral probability evolves similarly as the original distribution but with a drift lowered by the risk premium, λ (i.e.: it is normally distributed with mean $(\mu - \lambda)$ and volatility σ). It is important to notice that reducing the drift by the risk premium is not the same as assuming risk-neutrality, since returns still evolve as a random walk. Conceptually, the risk-neutral probability can be thought of as a distribution of returns that are stripped off the risk premium, so that the probability of a risky return within this distribution accounts simultaneously for the probability of the state and its risk.

4.7.2 Evolution of the Electricity Market Demand, Capacity and Price Indices

The evolution of a demand shock (x_{ij}) is modeled as a geometric Brownian motion with drift (Dixit and Pindyck, 1994), so that percentage changes in demand, $\Delta x_{ij}/x_{ij}$, are normally distributed. In this implementation we choose demand change to be perfectly correlated with the market return, and we model it as a random walk with a step, ε_{ij}, used to model the evolution of market returns.

$$\Delta x_{ij} = \alpha x_{ij} \Delta t + \sigma x_{ij} \sqrt{\Delta t} \varepsilon_{ij} = \alpha x_{ij} \frac{1}{n} + \sigma x_{ij} \frac{1}{\sqrt{n}} \varepsilon_{ij} \tag{4.24}$$

If changes in the demand shock are normally distributed, then the demand shock is *lognormally* distributed. Using Ito's lemma, it can be proved that, if $F_{ij} \equiv F(x_{ij}) = \log x_{ij}$,

the increments in the logarithm of demand evolve as a simple Brownian motion with drift:

$$\Delta F_{ij} = \left(\alpha - \frac{\sigma^2}{2}\right)\frac{1}{n} + \frac{\sigma}{\sqrt{n}}\varepsilon_{ij} \tag{4.25}$$

For a given time period, the demand shock can thus be expressed as a function of the demand shock in the previous period and the increment of the logarithm of demand:

$$x_{ij} = x_{i-1j}\exp\left(\Delta F_{ij}\right) \tag{4.26}$$

Figure 4-4 shows some results of the simulation. In the first panel, we show the logarithm of the market index (equal to the cumulative market return), following a Brownian motion with drift. The second panel shows the evolution of the demand for electricity. If the first two panels are compared, it becomes clear how these two variables are correlated. The third panel presents the evolution of electricity prices, which are connected with the demand through the price elasticity. Finally, the fourth panel presents the evolution of the installed capacity, which is a function of the electricity price and, as a result, it is also strongly correlated with demand.

4.7.3 Valuation of the Three Assets

Let's assume that the system is composed of multiple plants from two generating technologies: baseload and peaker, each of them with a different variable cost (c) and a different maximum output capacity (P). These are the technologies for which we want to price their risk through the financial options described above, assuming that plants only operate whenever the electricity price is above the variable cost of the plant (the strike price of the option) and that, when that occurs, they generate at their maximum output.

Therefore, the cash flows obtained by each call option will be the product of the spark spread (the difference between price and variable cost), and the hours that the plant is generating:

$$CF_{ij} = \max\left\{\rho_{ij} - c, 0\right\}\frac{8,760}{n} \tag{4.27}$$

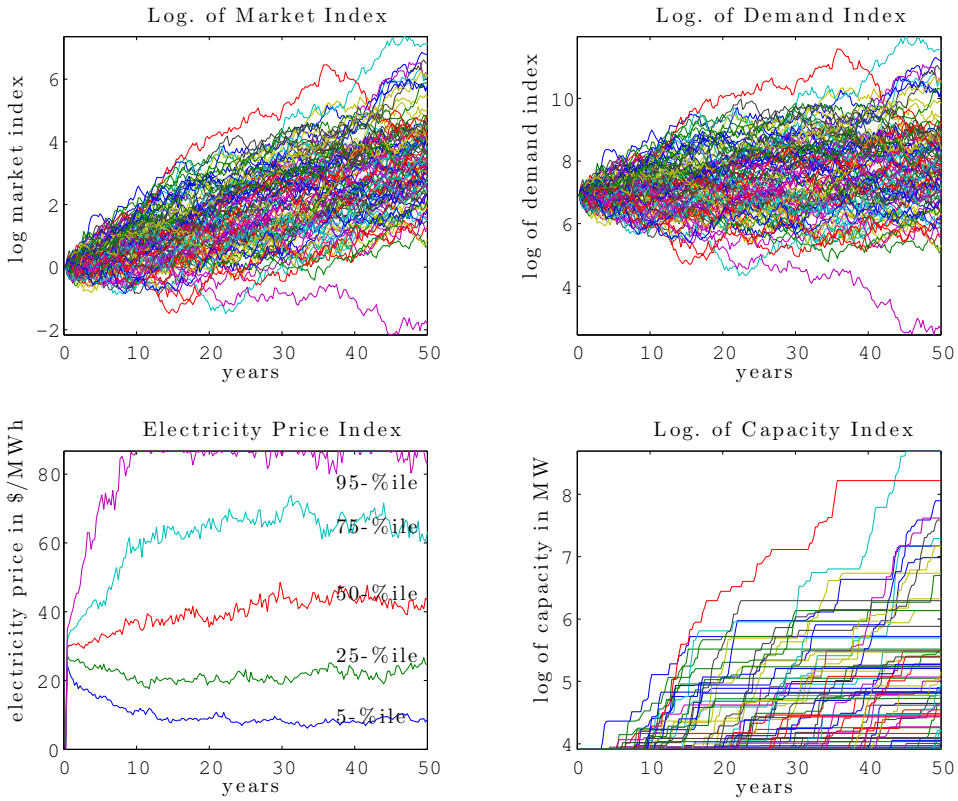


Figure 4-4: Market return, demand, electricity price and capacity results of a 50 year simulation.

Top left panel: natural logarithm of the market index evolution (cumulative market returns); top right panel: evolution of the log of the demand index; bottom left panel: 5, 25, 50, 75 and 95 percentiles of the price in [$/MWh]; bottom right panel: log of total installed capacity in the system in MW.

Other more complex and realistic models could have been used to calculate cash flows. However, this stylized model allows us to clearly reflect one particular phenomenon: the production from peaking units will only occur when prices are sufficiently high, making their cash flows more exposed to the market return than those of baseload units.

The asset can be valued as the sum of the present value of the cash flows received throughout all the periods considered:

$$PV = \sum_{i=0}^{N} \left(\sum_{j=0}^{J} CF_{ij} \pi_{ij} \phi_{ij} \right) \exp\left(-ri/n\right) = \sum_{i=0}^{N} \left(\sum_{j=0}^{J} CF_{ij} \pi_{ij}^{*} \right) \exp\left(-ri/n\right) \tag{4.28}$$

Table 4.5: Present value of the three assets for $x_{0j} = 20q_{0j}$.

PV_b [M$/MW]	PV_p [M$/MW]	PV_e [M$/MW]
3.2	0.0185	4.3

4.7.4 Implied Discount Rates for the Three Assets

If the expression above is compared with the analogous present value calculation using risk-adjusted discounting, one can see how it is possible to translate back to an average risk-adjusted discount rate, $\bar{\lambda}$:

$$PV = \sum_{i=0}^{N} \left(\sum_{j=0}^{J} CF_{ij}\pi_{ij} \right) \exp\left(-\bar{\lambda} i/n\right) \exp\left(-ri/n\right) = PV(\bar{\lambda}) \tag{4.29}$$

Having used the earlier expression to solve for the present value on the left-hand, the only unknown on the right-hand-side is the average risk-adjusted discount rate which we can then solve for. This equation, however, is hard to solve numerically, but it can be transformed it into an N-order polynomial if we apply the change of variable $s = \exp\left(-\bar{\lambda}/n\right)$. The roots of the resulting polynomial can be easily calculated with a commercial mathematical software like MATLAB:

$$\begin{aligned} PV(\bar{\lambda}) = PV(s) = \sum_{i=0}^{N} \left(\sum_{j=0}^{J} CF_{ij}\pi_{ij} \right) s^i \exp\left(-ri/n\right) = \\ = \mathbb{E}_j[CF_{0j}] + \mathbb{E}_j[CF_{1j}] \exp\left(-r/n\right)s + \ldots + \mathbb{E}_j[CF_{Nj}] \exp\left(-rN/n\right)s^N \end{aligned} \tag{4.30}$$

Impact of Initially Installed Capacity

In the particular example used in describing this implementation as well as in real life valuation exercises, initial conditions play an important role on the financial evolution of the asset. It is not the same starting with a generation deficit where prices are high and units are installed right away, as starting with a generation surplus where prices are low and units underutilized. In order to gain insight on how the initial conditions affect the

results, a sensitivity analysis is performed, calculating the technology-specific discount rates at different initial demand shock values, translating into an initial electricity price value.

Table 4.6: Impact of the initial electricity price, ρ_0, using 10 trials with 100 samples each.The standard deviation is given in parenthesis.

Initial demand shock/capacity	Initial electricity price [$/MWh]	$\bar{\lambda}_e$ [%]	$\bar{\lambda}_b$ [%]	$\bar{\lambda}_p$ [%]
5	7.27	1.89 (0.09)	3.32 (0.13)	5.33 (0.36)
10	12.70	1.66 (0.10)	2.49 (0.15)	4.86 (0.32)
20	28.58	1.26 (0.09)	1.62 (0.12)	3.90 (0.21)
30	44.99	1.00 (0.07)	1.24 (0.09)	3.71 (0.20)
40	61.96	0.86 (0.08)	1.04 (0.10)	3.49 (0.16)
50	79.56	0.78 (0.05)	0.94 (0.06)	3.37 (0.18)
60	86.71	0.69 (0.06)	0.83 (0.07)	3.05 (0.16)

These results show that lower ratios between the demand shock and generation (situations with a low electricity price) entail a larger risk for both baseline and peaker technologies since it will take longer to achieve higher prices. As the demand shock increases, so do prices, reducing the risk inherent to electricity prices. The difference in the technology-specific risk between baseload and peaking technologies is mostly due to the quantity risk, as the quantity generated by peakers is more exposed to the electricity price than baseload plants. The explanation for this is that units generate at times where price is greater or equal to their variable cost, and peaking units are more limited by this condition given their high variable cost. The effects described can be also observed in Figure 4-5 and Figure 4-6.

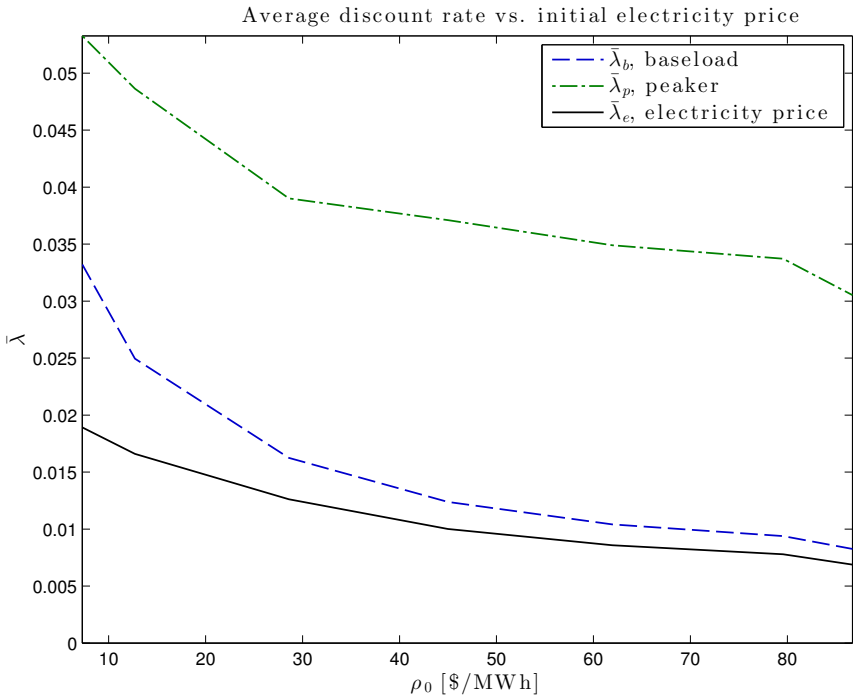


Figure 4-5: Average discount rate of the asset, $\bar{\lambda}$ vs. initial electricity price, ρ_0.

Plot of the average discount rate for the three assets considered, baseload ($\bar{\lambda}_b$), peaker ($\bar{\lambda}_p$), and financial option on the electricity price ($\bar{\lambda}_e$), as a function of the initial electricity price, ρ_0, which reflects the initial ratio between demand and capacity installed and the initial conditions upon which investment decisions will be made.

Robustness of the Solution

The robustness of the Monte Carlo simulation is tested by repeating the simulation with a different size of the number of scenarios sampled by period. The mean and the standard deviation of the distribution is approximated through bootstrapping ten samples of the estimated lambdas.

4.8 Conclusions

This chapter applied the stochastic discounting methodology, widely used in finance to value securities and assets, to the particular case of two electricity generation technologies. It presented the general algorithm with a numerical example and described the general mechanics

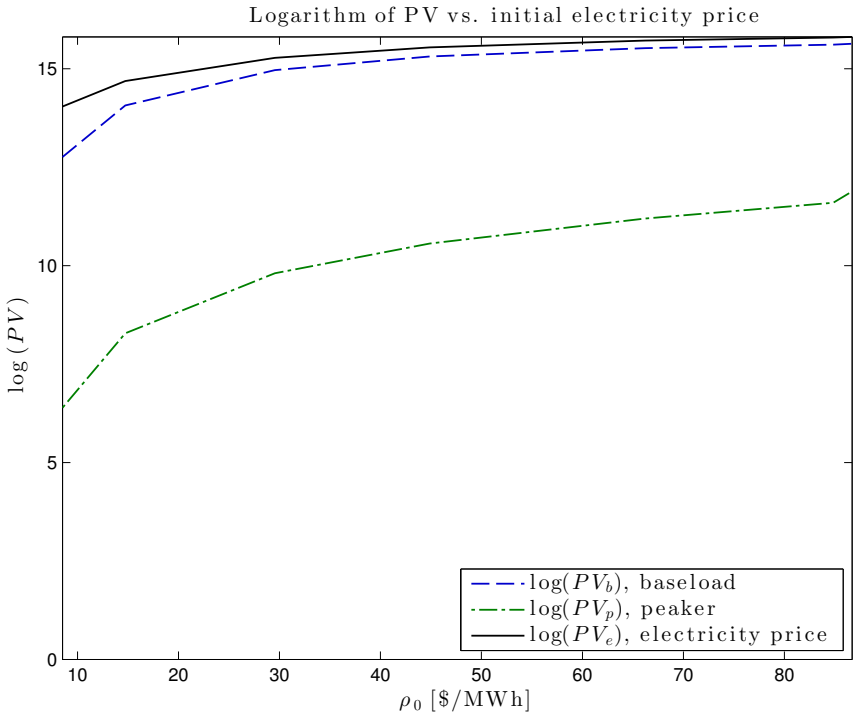


Figure 4-6: Log. of the present value of the asset, $\log(PV)$ vs. initial electricity price, ρ_0.

Plot of the average logarithm of the present value of the three assets considered, baseload ($\log(PV_b)$), peaker ($\log(PV_p)$), and financial option on the electricity price ($\log(PV_e)$), as a function of the initial electricity price, ρ_0, which reflects the initial ratio between demand and capacity installed and the initial conditions upon which investment decisions will be made.

of one particular implementation using Monte Carlo simulation. It then explained how the macro and the micro levels can be coupled with a model that forecasts cash flows contingent to the different realizations of an underlying risk factor.

It was demonstrated how different risk premia can be naturally derived from first principles for different electricity generation technologies, and how different ways to present risk can be used under different circumstances without changing the present value of the cash flows being discounted. It was shown that the risk attributable to each technology depends on the electricity price level at the time that the investment takes place. Finally, it was illustrated the advantages of stochastic discounting over risk-adjusted discounting –the commonplace in the electricity sector–, and how risk-adjusted discounting overestimates the

Table 4.7: Impact of the number of scenarios sampled by period during the simulation.The standard deviation is given in parenthesis.

Number of samples	$\bar{\lambda}_e$ [%]	$\bar{\lambda}_b$ [%]	$\bar{\lambda}_p$ [%]	time [s]
100	1.29 (0.06)	1.67 (0.08)	4.23 (0.33)	53
500	1.29 (0.03)	1.66 (0.04)	4.07 (0.17)	166
1,000	1.27 (0.02)	1.65 (0.03)	4.04 (0.10)	314
5,000	1.27 (0.01)	1.64 (0.02)	4.04 (0.04)	1,534

value of cash flows that are more exposed to the risk factor and underestimates the value of cash flows less exposed.

Stochastic discounting can serve as an extremely helpful tool to make better investment decisions in power generation, as long as we are capable to establish sound relationships between the underlying risk factor and the assets cash flows.

Chapter 5

Conclusions

5.1 Summary

This thesis has shown the market impacts of intermittent renewables (e.g., wind and solar photovoltaic) on future investments in electricity markets, focusing on the average profit made by each generation technology and the risks they are exposed to. The analysis distinguished between two situations: a system where the generation capacity mix is optimized ('adapted') from 'greenfield', given a certain amount of installed renewable capacity; and a system optimized for zero renewables, but where, on top of this optimal mix, some amounts of renewable generation have been added ('non-adapted'). Concurrently with distinguishing between these two situations, this thesis explored how a set of wholesale market pricing rules representative of the diversity of existing regulatory frameworks impact on the recovery of the total generation costs. Accordingly, in the two adaptation cases explored, generators could be at risk of not recovering their costs: in the 'adapted' situations, generators are exposed to the risk of pricing rules, despite having a well-adapted generation mix; in the 'non-adapted situation, besides the pricing rule risk, generators are exposed to the risk of new regulation that unexpectedly brings extra renewables on top of a well-adapted generation mix.

Besides looking at the average profit of each generation technology considered, the cases studied were also assessed from a multi-dimensional perspective looking at the capacity

installed from each technology, capacity factor, energy contribution, startup regime, total system emissions, renewable curtailment, and the average cost paid by consumers.

Moreover, this thesis proposed a valuation method that dynamically prices the risk seen by investors in generation, such as the regulatory risk from unexpectedly deploying renewable generation. Overall, the thesis provides insights about any additional risks for future investors that a strong presence of renewables might introduce in wholesale markets, and the role that the design of market rules would play in this. Each chapter in this thesis included at the end a detailed conclusions section, and in this chapter we summarize the most important ones.

5.1.1 Modeling Electricity Markets with Renewables

This thesis analyzed the market impacts of renewables using IMRES, a capacity expansion model with unit commitment constraints, inspired in previous works that used a similar approach to study the impact of operational flexibility needs on capacity decisions, in the context of power systems with renewables. IMRES was designed to include the extra cost from cycling the plants in investment decisions. Specifically, IMRES accounts for the extra cost derived from respecting the operational constraints of the plants, and the chronological variability of the net load. In addition to these features, IMRES is unique in that investment decisions are taken at the individual power plant level, enabling a detailed cost recovery analysis of every power plant in the system. With IMRES, as with any other generation capacity expansion model, in equilibrium and in the absence of internal or common capacity constraints on specific technologies, the profits obtained by all power plants in the system are close to zero, with a small unavoidable divergence originated in the lumpiness introduced by the discrete size of the plants and the differences in pricing rules originated by the diverse treatments given to the non convex costs, such as those from starting up the plants.

One of the biggest challenges from using formulations that combine capacity decisions with unit commitment is the high dimensionality of the resulting problem. Formulations that do not consider detailed commitment constraints overcome this challenge using decom-

position techniques such as Benders'. However, formulations like IMRES where the solution space of the commitment subproblem is nonconvex, cannot guarantee the convergence of the solver to an optimal solution because of the duality gap. Instead, IMRES simplifies the problem using a limited number of representative weeks in the unit commitment component of the model. This simplification reduces the dimensionality of the problem, while retaining the chronological variability of the net load.

The week selection algorithms proposed by some authors revolve around choosing the weeks containing extreme values of demand or wind output –combinations of high and low values of both–, to guarantee that these events are accounted for in the resulting capacity mix. Yet, the number of extreme events with more than one renewable energy resource (e.g., wind and solar PV) increases significantly the number of representative weeks needed to characterize both the load and the renewable resource in the system –think about the number of different combinations of high and low demand, and high and low values of wind and solar output–. Instead, the approach used with IMRES focuses directly on characterizing the net load duration curve in the system. By focusing on the net load, we simultaneously take into account demand and renewable energy output, as well as the correlation between them. The algorithm is based on a simple idea: the weeks selected are the ones that combined minimize the error between the actual net load duration curve and an approximation that is built with the weeks selected. This algorithm, however, does not explicitly select the weeks on the basis of representing the chronological variability of the net load, and the weeks selected with this algorithm need to be tested on how well they represent this aspect. Still, compared to other week selection methods, the algorithm proposed with IMRES performs better in terms of representing the net load in the system and the chronological variability.

5.1.2 Electricity Market Impacts with and Without Adaptation

Throughout the analyses we distinguished between two situations where conventional generators could be at risk of not recovering their costs. In the first situation, generators are exposed to the risk of the pricing rules, despite having a well-adapted generation mix. In the

second situation, generators are exposed to pricing rules and to new regulation unexpectedly bringing extra renewables on top of a well-adapted generation mix. In reality, most power systems will be somewhere in between these two situations. In particular, as renewable capacity incentives are frequently administered without allowing other players in the market to anticipate, power systems will be further away from equilibrium.

Profit Results

The metric used to represent the cost recovery of the different technologies in the system is the average profit obtained by plants from the same technology (the difference between revenues and total cost, expressed in terms of the total cost). Profits, in turn, depend on how the optimal mix evolves in the future with more renewables, and whether new renewable deployment occurs expectedly or not. We summarize first the results obtained for the resulting capacity mix, capacity factor and wholesale prices in order to better explain the main results on profits, summarized subsequently.

With adaptation, results showed that capacity evolves towards a more flexible mix, that can better complement renewable output in order to instantaneously balance generation and demand. Specifically, nuclear and coal capacities are progressively displaced, introducing more gas-fired power plants (CCGTs and OCGTs). The average capacity factor –the level of utilization of the plants– with adaptation remains relatively stable for all the technologies across the different renewable scenarios studied (98% for nuclear, 92% for coal, 45% for CCGTs and 8% for OCGTs). Without adaptation, the thermal capacity mix remains constant for all renewable scenarios, but the effect on the capacity factor considerably changes depending on the technology: the utilization of nuclear remains relatively stable at 95%, while the utilization of coal, CCGTs and OCGTs decreases as these technologies get displaced by zero-variable-cost renewable energy. Specifically, between $S = 0$ and $S = 10$, the utilization of coal decreases from 92% to 30%, the utilization of CCGTs decreases from 45% to 10%, and the utilization of OCGTs decreases from 8% to 5% in $S = 1$ and remains at this level with subsequent larger penetration scenarios.

The different adaptation situations have also a strong impact on the average prices obtained. We took the dual variable of the demand balance equation as a proxy of the hourly electricity price, and we observed that under adaptation prices remain relatively stable on average across the renewable scenarios studied. Conversely, under non-adaptation, prices decreased as a larger renewable capacity was installed in the system. This later effect is caused by two main factors: the first one is that renewables have a zero variable cost and push the most expensive technologies out of the merit order; the second is that renewables are added on top of a system that is already adequately sized, creating an excess of capacity situation and bringing the hours with non-served energy close to zero.

Finally, the main results on profits tied together the effects on wholesale prices, the utilization of the plants, and their cycling regime, reflecting the risk that generation technologies see under the two situations considered. With adaptation, the average prices and the capacity factor of the plants are shown to be stable for different levels of penetration of renewables, which result in profits that are close to the theoretical market equilibrium result. Conversely, without adaptation, we saw that technologies are exposed to lower prices[1] , a reduction of the capacity factor and an increase of the cycling regime. Yet, generating technologies in the non-adapted case were not exposed equally to all these three effects: we found that nuclear plants are mostly exposed to the 'price effect' without having their capacity factor significantly changed; while coal plants, CCGTs and OCGTs are exposed to the 'price effect', and a reduction of the capacity factor –'quantity effect'–, making plants incur in losses. This result underlines the importance of having an adapted system to allow all plants to recover costs.

Additional System-wide Results

Besides results on profits, we also obtained other side information that could be valuable in order to gain more insight on other system-wide considerations, like the use of different fuels and the emissions associated with generation.

[1]It is important to remind here that each technology sees a different price to the other, as the hours during which different plants are dispatched are different.

The varying capacity factor of plants results in a different energy contribution of the existing technologies to the total electricity demand. With adaptation, the share of nuclear power steadily declines (from 70% without renewables to 32% with the highest renewable penetration level considered), and the contribution from CCGTs increases as more renewables are installed in the system. Without adaptation, nuclear provides a steady 70% of the total electricity demand across all renewable scenarios because of its high minimum stable output and its high startup cost, the contribution of coal is around 5%, and the remaining share is supplied by CCGTs, which declines as more renewables are introduced in the system (from 23% without renewables to 5% with the highest renewable penetration level considered) .

The way in which electricity is generated reflects directly on the total emissions produced by the system. With adaptation, as renewables are introduced, emissions drop sharply to 60% of the total emissions with no renewables, because coal is the first technology being displaced. However, with larger renewable shares, we observe that total emissions hit a floor corresponding to the emissions from gas-fired power plants, because once a certain renewables penetration level has been reached, electricity is solely produced by a blend of renewables, gas and nuclear power. Without adaptation, emissions decline slower, but they can be reduced past the floor encountered with adaptation. The reason for this is that without adaptation, electricity is mostly generated by a combination of renewables and nuclear, with curtailment providing most of the flexibility needed to balance generation and demand. It is expected that if more ambitious emissions targets are to be reached in equilibrium, the system will require the deployment of other elements such as storage and demand response, that increase the utilization of renewables as well as the flexibility in the system, while reducing the cycling from conventional thermal generators (mostly provided by gas-fired power plants, as it resulted in the simulations with the adapted system).

5.1.3 Impact of Bidding Rules

One of the initial hypotheses of this research is that greater renewable shares increase the cycling regime of thermal power plants. We observed that this effect actually takes place, although the technologies that see their number of startups increased differ depending on the adaptation situation. With adaptation, coal power plants are phased-out and the number of cycles performed is reduced as more renewables are installed in the system; the average number of start ups per CCGT varies in each scenario and does not show any clear trend; the cycling regime of OCGTs, however, increase as more renewables are introduced in the system. Without adaptation, the situation is remarkably different as now it is coal power plants that bear the extra cycling, with CCGTs and OCGTs not showing a clear trend upwards or downwards. The tendency is more clear if we look at results in terms of the total aggregated number of startups performed by individual technologies: overall, with adaptation the extra cycling falls on gas-fired power plants, while, without adaptation, the extra cycling falls on coal plants.

These startup costs are nonconvex costs that are not reflected in the dual variable of the demand balance equation (they are not reflected in the electricity price). Therefore, having an adapted system, does not guarantee the cost recovery of the plants, as startup costs must also be included somehow in the market price in order to recover total costs (operating and investment), in order to have an adequate capacity mix in equilibrium. These nonconvex costs are typically associated with the costs of starting up and the lumpiness introduced by the indivisibility of plants. Besides these two factors common to all power systems, there is an additional source of discontinuity in the cost function in IMRES originated in the selection of a small number of weeks to represent a full year net load series (e.g., given a one year net load series with a total of 5 hours with non-served energy, if one of the 5 hours with non-served energy appears in the four weeks selected, as we expand by a factor of 13 the hours in the four weeks selected to build the NLDC approximation, the NLDC approximation will appear with 13 hours of non-served energy as opposed to 5). The effects produced by this source of lumpiness were mitigated by selecting a relatively low value of lost load, reducing

the effect of non-served energy on plants' profits.

We tested three different bidding mechanisms aimed at guaranteeing that plants recover nonconvex operating costs, in addition to a nave 'do-nothing' case to assess the cost that is not recovered taking the dual variable of the demand balance constraint as hourly price. [2]

Results showed that, besides helping plants to recoup operating costs, the three bidding mechanisms studied have different implications on the total cost recovery of plants and therefore on the cost paid by consumers for electricity. First, we found that a side payment, equal to the operating costs that are not recovered through energy sales, brings all technologies closer to full cost-recovery at a low cost to the consumer. Second, if the minimum cost uplift that guarantees that all plants recover operating costs is added on top of the hourly price, the resulting profits are greater than with a side payment, as the final price, including the uplift, affects all units. We also found that the uplift contributes towards recovering investment costs, and that all technologies were able to recover total costs on average. Yet, the cost to the consumers is also higher in the short-term. In the long-term, having a higher price might attract new entrants and, in turn, reduce the price. Lastly, when plants internalize their startup costs in a simple bid, we found that prices as well as profits were significantly higher than in all previous cases, overcompensating all technology classes. These positive profits are an incentive for investors to install more capacity, which results in an equilibrium point with more generating capacity installed than it is actually needed (as given by the optimal capacity mix determined by IMRES). We proposed a heuristic method to iterate on the solution until a new equilibrium where no new entrants would make a profit is found[3]. After applying the heuristic algorithm, and with an equilibrium solution, we observed that prices and profits were still and higher than with other bidding rules, imposing an unnecessary cost burden on the consumer.

[2]In all the bidding rule cases studied, the prices were formed starting from the optimal mix and optimal dispatch that were obtained by centralized optimization taking all costs into account (i.e., the capacity and the dispatch in all cases is the same).

[3]Note, however, that with simple bids the 'adapted' generation mix and the unit commitment in reality would have been different to the ones resulting from centralized optimization, as the commitment algorithm is solely based on the simple bids received by the market operator. Yet, we started from the same generation mix derived from using a centralized unit commitment, in order to be able to isolate the effect of the bidding rules studied on prices from other effects.

Finally, when the differences in profits under adaptation resulting from the three different bidding rules studied were compared to those under non-adaptation, we saw that the magnitude of the differences in profit from the three ways of accounting for nonconvex costs were significantly smaller than the profit gaps caused by the excess of capacity in the non-adapted cases. This result highlights that the most important factor affecting the profits of plants is how far the capacity mix is from equilibrium.

5.1.4 Regulatory Implications

Unanticipated deployment of renewable generation has been shown to produce economic losses to some generators[4]. These losses are the result of regulatory shocks that, given their unanticipated character, disrupt the existing market equilibrium, making some of the incumbent investments redundant. The losses can be thought of as *stranded costs*, using the same denomination for the economic losses induced during the transition from the traditional vertically integrated utility business model to competitive markets in generation.

In the short-term, the reduction of the capacity factor of any given plant can lead to two outcomes: the owner of the plant might decide to keep the plant operating in order to minimize its economic losses; or, in the unlikely situation that the total (fixed and variable) operating costs are greater than the revenues that the plant receives, they might opt to mothball or close the plant, which could have an impact on the security of supply of the system as thermal capacity decreases.

In the long-term, regulatory uncertainty increases the risk of the plant. If this uncertainty is found to have a correlation with macroeconomic factors, then the risk associated with that uncertainty cannot be diversified away in a portfolio with other different assets and has to be priced. This increase of risk reduces the value of thermal generation investments, making them less attractive compared to other assets available in the financial market. Ultimately, increasing the regulatory uncertainty seen by power plants by not allowing for anticipation as

[4]In this thesis we only explore the day-ahead market, and all technologies are found to incur losses under non-adaptation. However, in reality, as renewable capacity increases, so does the need of reserves, and some plants that see their utilization go down, might choose to participate in reserves markets, possibly mitigating this negative impact. The study of this effect is left for future research.

new renewable capacity is introduced in the system undermines the attractiveness of thermal plant investments, as well as the security of supply in the system.

In the face of the negative effects from non-adaptation, there are two courses of action that the regulator could take ex-post the regulatory shock: let the generators bear with the economic losses from the unexpected new renewable capacity, or compensate generators for these losses. If the regulator chooses not to do anything, the short-term cost to the consumer is lower, but in the long-term investors in generation would see a high regulatory risk in the market, investment will be suboptimal and market prices will increase. Conversely, the regulator may decide to recognize the losses and compensate generators for them at the expense of the final consumer. This option would reduce the regulatory risk seen by investors, but would impose a temporary burden on the consumer.

Regarding bidding rules, we found that the three rules studied guarantee the recovery of operating cost, as it was expected. However, these rules are not designed to recover operating costs, but total costs. Therefore, we then explored the different tradeoffs among them, focusing on how each affected the recovery of all remaining costs (fixed operating costs and investment costs) for the inframarginal units, and how that translated into different costs to the consumers. We found that providing generators with a side payment for the operating costs that are not recovered through energy sales does not affect the recovery of investment costs. Conversely, we found that adding to the price an uplift had an impact of the recovery of investment costs and, on average, we observed that all technologies recovered total costs. The price uplift rule increased the cost to the consumer compared to the cost with a side payment, but not by a great amount. Lastly, we found that simple bids also allow all technologies to recover total costs but translate into an excessive cost to consumers[5].

[5]This conclusion, however, is subject to the caveat indicated above, that we are not starting from the generation mix that would be obtained under the simple bids market rule, as the commitment results from settling the simple bids and not from a centralized unit commitment.

5.1.5 Accounting for Risk in Generation Investments

In general, the risk seen by power plants is very nonlinear, as an increase in the market return of a portfolio representing all the assets in the market does not translate in a proportional increase in the returns obtained by the power plants investment. Moreover, the distribution of uncertainty across time is very heterogeneous, as the risk associated with some factors –like regulatory shocks– does not increase proportionally with time. These two realities make power plant investments to separate from the assumptions with traditional ways of pricing risk like the CAPM. In this thesis we proposed stochastic discounting as an alternative to the CAPM to price risk more accurately. Additionally, we proposed an algorithm to couple the stochastic discounting method with Monte Carlo simulation, improving the capability of the method to account for uncertainty.

Using a simple stylized capacity expansion model, we analyzed the intrinsic risk of two technologies –baseload and peaker–, and we found similar effects to the ones found in the first part of this thesis: baseload plants and peakers are exposed to a 'price effect', while peaking plants are additionally exposed to a 'quantity effect', as their output is reduced when prices are low. Therefore, we found that peaking technologies bear a higher risk than baseload, debunking the notion that nuclear plants have a higher risk, arguing that they have a higher capital cost. This notion is wrong because, although it is true that nuclear plants have a higher capital cost, they can recover these costs easier than a peaker plant, as peakers are more exposed to the electricity price risk factor.

5.2 Contributions

This research assessed the impacts of large shares of intermittent renewables on the risk seen by thermal generators of not recovering their costs, focusing on two main risk factors: the risk of the pricing rules used to recover operating costs, and the risk of having unexpected additions of renewables capacity on top of a well-adapted generation mix. Additionally, this thesis proposed a method to accurately price the risk of investments in electricity generation.

In particular, the contributions of this thesis are:

1. *Assessing the Extra Risk Brought in by Intermittent Renewables*: producing a comprehensive cost-recovery analysis of power plants, quantifying the profits obtained under three different bidding rules, and under two different adaptation situations. This thesis showed that the regulatory risk from adding new unexpected renewable capacity to the system can cause significant losses to a generation mix optimized for a different renewable capacity level. The thesis also showed that in equilibrium, thermal plants are also exposed to the differences in remuneration brought by the three bidding rules studied. However, it was also shown that this risk is smaller than the risk of unexpectedly adding extra renewable capacity to an adapted generation mix.

   IMRES provides a versatile test-bench for a number of experiments and, besides the analyses conducted for this thesis, IMRES has been used to assess the economic benefit of three different flexibility options (demand side management, storage and 'flexible' plants) in power systems with a large penetration of intermittent renewables (IEA, 2014).

2. *Modeling Power Systems with Renewables*: introducing a new formulation for analyzing electricity markets based on a capacity expansion formulation with unit commitment constraints. The formulation is new in that it simultaneously reflects the impact of operation constraints on cost, accounts for the chronological variability of the net load, and considers building decisions on an individual power plants basis. Additionally, this thesis has shown that four weeks selected based on approximating the NLDC, could also provide a good approximation of the full-year series with regard to commitment results. For each application of the selection week method, three error metrics have been suggested to test that this result actually holds. Moreover, this thesis presented a heuristic method that iterates on possible capacity expansion solutions until equilibrium is reached.

   IMRES provides a versatile test-bench for a number of experiments and, besides the analyses conducted for this thesis, IMRES has been used to assess the economic benefit

of three different flexibility options (demand side management, storage and 'flexible' plants) in power systems with a large penetration of intermittent renewables (IEA, 2014).

3. *Evaluation of Risk*: bridging the gap between the methods currently used in the valuation of power plant investments and modern financial tools, by making explicit the relationship between the cash flows obtained by two different generating technologies and the evolution of an index representing the average market returns in the economy. Additionally, this thesis proposed an algorithm to price risk dynamically, exploiting the capabilities of the Monte Carlo method to consider many uncertainty paths.

5.3 Further Research

This thesis has increased our understanding of how intermittent renewables impact power systems and electricity markets, focusing on the day-ahead market. At the same time, this thesis opens the floor for new research that could improve or complement some of the methods and experiments presented. Following the same structure of the previous section, we next describe possible future research that could overcome some of the limitations in the works presented in this thesis:

1. *Assessing the Extra Risk Brought in by Intermittent Renewables*: the analyses in this thesis were limited to the day-ahead market. One important improvement would be studying the interaction between renewables, reserve requirements and reserve markets, and how reserve markets might help to reduce the profit gap under non-adaptation. IMRES was prepared to study this interaction, and reserves were included in the main formulation of the model. However, studying the market interaction between the day-ahead market, reserve markets and how that affects investment decisions required using the dual values of the different reserve constraints. During the experimentation in this thesis, whenever these values were called on the model, the solver would not produce a solution. Further work could be used in calibrating IMRES so that it provides the

dual values of the reserve constraints and allows the study of the interaction between day-ahead and reserve markets.

Additionally, further work could study the market effects of dynamic elements such as storage or demand response in systems with renewables. As these elements become more important to reduce renewable curtailment and integrate large shares of intermittent renewables, studying their interaction with electricity markets will become increasingly relevant.

In the same line, it would be interesting to explore the optimal capacity of storage or demand side response necessary to complement power systems with renewables under different situations: a power system with 100% renewable shares, a power system with renewables and nuclear power, a power system with renewables and gas, etc.

2. *Modeling Power Systems with Renewables*: IMRES' unit commitment module used a week selection method that is solely based on approximating the NLDC. However, although the commitment solution with the approximation has been very close to the solution obtained with a unit commitment of the full-year series for the cases studied in this thesis, this result is not mathematically guaranteed by the selection method. One possible improvement could be considering the chronological variability of the net load as an additional criterion for selecting representative weeks, to guarantee that variability is well-represented in the four week approximation.

   Alternatively to models based on using a small set of representative weeks, there are other capacity expansion models with unit commitment constraints that consider the full-year in the commitment component of the model. Based on this other approach, one possible line of research would be analyzing the market effects of renewables on power systems as it was done in this thesis. Using the full-year data in the commitment instead of representative weeks would increase the accuracy of the analyses. Yet, simplifications in the profit calculations, considering technologies instead of individual plants as it is done here, would likely be required.

Future research on capacity expansion models that account for flexibility needs, could address other modeling techniques, such as characterizing unit commitment as a state space, and incorporate some of the operating constraints that are not included at present, and embed it into a capacity expansion formulation.

IMRES treats the electricity network as a 'copperplate', ignoring the effect of transmission constraints and losses. In some power systems, transmission constraints are found to be the first cause for curtailment. Therefore, one important improvement to this model would be adding transmission constraints, while keeping solution time within reasonable limits.

Another assumption made with IMRES is that the market is perfectly competitive, which is valid for systems where the market share of the individual agents is relatively low. However, many power systems still operate under oligopolistic conditions, with some large utilities owning generation assets of multiple technologies and abusing of market power. Analyzing the combined effect of oligopolistic practices with renewables could provide a more accurate picture of the market effects in this type of systems, and could help explain the bidding behavior of some agents.

3. *Evaluation of Risk*: the microeconomic model of the power industry used in the valuation of generating technologies is based on the Leahy model, which bears little resemblance to the dispatch models used in reality to determine electricity prices. Further work should adapt this model to incorporate some of the features of real dispatch models to increase the accuracy with which cash flows are calculated.

Bibliography

AIP (2007). *The Bidding Code of Practice. A Response and Decision Paper.* All Island Project, AIP-SEM-07-430,30.

Baldick, R. (1995). The generalized unit commitment problem. *IEEE Transactions on Power Systems*, 10(1):465–475.

Baritaud, M. (2012). *Securing Power During the Transition. Generation Investment and Operation Issues in Electricity Markets with Low Carbon Policies.* International Energy Agency. Insight Series, Paris.

Batlle, C. and Perez-Arriaga, I. J. (2008). Design criteria for implementing a capacity mechanism in deregulated electricity markets. special issue on capacity mechanisms in imperfect electricity markets. *Utilities Policy*, 16(3):184–193.

Batlle, C. and Rodilla, P. (2013). An enhanced screening curves method for considering thermal cycling operation costs in generation expansion planning. *IEEE Transactions on Power Systems*, 28(4):3683–3691.

Bertsimas, D. and Tsitsiklis, J. (1997). *Introduction to Linear Optimization.* Athena Scientific.

Bertsimas, D. and Weismantel, R. (2005). *Optimization Over Integers.* Dynamic Ideas, Belmont, MA.

Birge, J. R. and Louveaux, F. (1997). *Introduction to Stochastic Programming.* Springer, New York.

Brealey, R. A., Myers, S., and Allen, F. (2008). *Principles of Corporate Finance.* Mc Graw-Hill Irwin, 9th edition.

Brown, P. (2012). *U.S. Renewable Electricity: How Does Wind Generation Impact Competitive Power Markets?* Congressional Research Service. Report for Congress 7-5700 R42818.

Bushnell, J. (2010). *Building Blocks: Investment in Renewable and Non-Renewable Technologies.* Energy Institute at Haas, Berkeley, CA.

California ISO (2011). Flexible ramping constraint. Technical bulletin, 2011-02-01, California ISO.

Cerisola, S., Baillo, A., Fernandez-Lopez, J., Ramos, A., and Gollmer, R. (2009). Stochastic power generation unit commitment in electricity markets: A novel formulation and a comparison of solution methods. *Operations Research*, 57(1):32–46.

Cochrane, J. H. (2005). *Asset Pricing: Revised Edition*. Princeton University Press.

Copeland, T., Weston, J. F., and Shastri, K. (2005). *Financial Theory and Corporate Policy*. Pearson Education, fourth edition.

Cox, J. C., Ross, S. A., and Rubinstein, M. (1979). Option pricing: A simplified approach. *Journal of Financial Economics*, 7(1979):229–263.

Cox, J. C. and Rubinstein, M. (1985). *Options Markets*. Prentice-Hall.

de Jonghe, C., Delarue, E., Belmans, R., and D'haeseleer, W. (2011). Determining optimal electricity technology mix with high level of wind power penetration. *Applied Energy*, 88(6):2231–2238.

de Sisternes, F. J. (2011). Quantifying the combined impact of wind and solar power penetration on the optimal generation mix and thermal power plant cycling. In *Young Energy Economists and Engineers*, Madrid, Spain.

de Sisternes, F. J. (2013). Investment model for renewable electricity systems (imres): an electricity generation capacity expansion formulation with unit commitment constraints. *CEEPR Working Paper Series*, (16).

de Sisternes, F. J. and Parsons, J. E. (2014). A dynamic model for risk pricing in generation investments. preliminary draft.

de Sisternes, F. J. and Webster, M. D. (2013). Optimal selection of sample weeks for approximating the net load in generation planning problems. *ESD Working Paper Series*, (03).

de Sisternes, F. J., Webster, M. D., and Perez-Arriaga, I. J. (2014). The impact of bidding rules on electricity markets with intermittent renewables. *IEEE Transactions on Power Systems*.

Delarue, E., Luickx, P., and D'haeseleer, W. (2007). The actual effect of wind power on overall electricity generation costs and co2 emissions. *Energy Conversion and Management*, 50:1450–1456.

DENA (2010). Dena grid study ii. integration of renewable energy sources in the german power supply system from 2015-2020 with an outlook to 2025. Technical report, DENA - German Energy Agency.

Denholm, P., Ela, E., Kirby, B., and Milligan, M. (2010). *The Role of Energy Storage with Renewable Electricity Generation*. National Renewable Energy Laboratory (NREL).

Denholm, P. and Margolis, R. M. (2007). Evaluating the limits of solar photovoltaics (pv) in electric pwoer systems utilizing energy storage and other enabling technologies. *Energy Policy*, 35:4424–4433.

Dilon, T. S., Edwin, K. W., Kochs, H. D., and Taud, R. J. (1978). Integer programming approach to the problem of optimal unit commitment with probabilistic reserve determination. *IEEE Transactions on Power Systems*, PAS-97(6):2154–2166.

Dixit, A. K. and Pindyck, R. S. (1994). *Investment Under Uncertainty*. Princeton.

Duffie, D. (2001). *Dynamic Asset Pricing Theory*. Princeton University Press, Princeton, NJ.

Ehrenmann, A. and Smeers, Y. (2010). *Stochastic Equilibrium Models for Generation Capacity Expansion. Stochastic Optimization Methods in Finance and Energy: New Financial Products and Energy Market Strategies*. Springer.

EPRI (1980). *EGEAS: Electric Generation Expansion Analysis System, version 9.02B*. Electric Power Research Institute (EPRI).

Fabrizio, K. R., Rose, N. L., and Wolfram, C. D. (2007). *Do Markets Reduce Costs? Assessing the Impact of Regulatory Restructuring on US Electric Generation Efficiency*.

Fink, S., Mudd, C., Porter, K., and Morgenstern, B. (2009). *Wind Energy Curtailment Case Studies: May 2008 - May 2009*. National Renewable Energy Laboratory (NREL).

Floudas, C. A. (1995). *Nonlinear and Mixed-Integer Optimization: Fundamentals and Applications*. Oxford University.

GAMS (2014). Cplex12. Technical report.

Geoffrion, A. M. (1972). Generalized benders decomposition. *Journal of Optimization Theory and Applications*, 10(4):237–260.

Green, R. (2000). Competition in generation: The economic foundations. *Proceedings of the IEEE*, 88:128–139.

Gribik, P. R., Hogan, W. W., and Pope, S. L. (2007). *Market-Clearing Electricity Prices and Energy Uplift*.

Hasche, B. (2010). General statistics of geographically dispersed wind power. *Wind Power*, 13:773–784.

Hedman, K. W., O'Neill, R. P., and Oren, S. S. (2009). Analyzing valid inequalities of the generation unit commitment problem. *Power Systems Conference and Expostion*, pages 1–6.

Hobbs, B. (1995). Optimization methods for electric utility resource planning. *European Journal of Operations Research*, 83:1–20.

Hogan, W. W. (2005). On an "energy-only" electricity market design for resource adequacy. Technical report, Center for Business and Government, John F. Kennedy School of Government, Harvard University, Cambridge, MA.

Holttinen, H., Kiviluoma, J., Estanqueiro, A., Gomez-Lazaro, E., Rawn, B., Dobschinski, J., Meibom, P., Lannoye, E., Aigner, T., Wan, Y. H., and Milligan, M. (2010). *Variability of Load and Net Load in Case of Large Scale Distributed Wind Power*.

Holttinen, H., Meibom, P., Orths, A., O'Malley, M., Ummels, B. C., and Tande, J. O. (2008). Impacts of large amounts of wind power on design and operation of power systems. Technical report, Results of IEA Collaboration.

IEA (2011). Harnessing variable renewables. a guide to the balancing challenge. Technical report, International Energy Agency (IEA), Paris.

IEA (2013). *Medium-Term Renewable Energy Market Report 2013*. International Energy Agency, Paris.

IEA (2014). *The Power of Transformation. Wind, Sun and the Economics of Flexible Power Systems*. International Energy Agency, Paris.

Ingersoll, J. E. (1987). *Theory of Financial Decision Making*. Studies in Financial Economics. Rowman and Littlefield.

Joskow, P. L. and Schmalensee, R. (1983). *Markets for Power*. MIT Press.

Kahn, A. (1971). *The Economics of Regulation. Principles and Institutions*. MIT Press.

Kirschen, D. S., Ma, J., Silva, V., and Belhomme, R. (2011). Optimizing the flexibility of a portfolio of generating plants to deal with wind generation. In IEEE, editor, *IEEE Power and Energy Society (PES)*, Detroit, MI.

Kirschen, D. S. and Strbac, G. (2004). *Fundamentals of Power System Economics*. Wiley.

Lannoye, E., Flynn, D., and O'Malley, M. (2011). The role of power system flexibility in generation planning. Detroit, MI. IEEE Power and Energy Society.

Lannoye, E., Milligan, M., Adams, J., Tuohy, A., Chandler, H., and Flynn, D. (2010). Integration of variable generation: Capacity value and evaluation of flexibility. In IEEE, editor, *IEEE Power and Energy Society*.

Leahy, J. V. (1993). Investment in competitive equilibrium: The optimality of myopic behavior. *Quarterly Journal of Economics*, 108(4-11):1105–33.

Li, X., Tomasgard, A., and Barton, P. I. (2011). *Nonconvex Generalized Benders Decompostion for Stochastic Separable Mixed-Integer Nonlinear Programs*. MIT Department of Chemical Engineering, Cambridge, MA.

Ma, J., Silva, V., Belhomme, R., Kirschen, D. S., and Ochoa, L. (2013). Evaluating and planning flexibility in sustainable power systems. *IEEE Transactions on Sustainable Energy*, 4(1):200–209.

McKay, M., Beckman, R. J., and Conover, W. (1979). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 21(2):239–245.

Milligan, M., Donohoo, P., Lew, D., Ela, E., Kirby, B., and Holttinen, H. (2010). *Operating Reserves and Wind Power Integration: An International Comparison*. National Renewable Energy Laboratory (NREL).

MIT (2011). *The Future of Natural Gas*. MIT Energy Initiative, Cambridge, MA.

Morales, J. M., Conejo, A. J., and Perez-Ruiz, J. (2009). Economic valuation of reserves in power systems with high penetration of wind power. *IEEE Transactions on Power Systems*, 24(2):900–910.

Myers, S. C. and Howe, C. D. (1997). *A Life-Cycle Financial Model of Pharmaceutical R&D*. Massachusetts Institute of Technology, Sloan School of Management, Cambridge, MA.

Newell, S. (2012). *ERCOT Investment Incentives and Resource Adequacy*. The Brattle Group.

O'Neill, R. P., Sotkiewicz, P. M., Hobbs, B. F., Rothkopf, M. H., and Jr, W. R. S. (2005). Efficient market-clearing prices in markets with nonconvexities. *European Journal of Operations Reseach*, 164(2005):269–285.

Ostrowski, J., Anjos, M. F., and Vannelli, A. (2012). Tight mixed integer linear programming formulations for the unit commitment problem. *IEEE Transactions on Power Systems*, 27(1):39–46.

Palmintier, B. (2012). *Incorporating Operational Flexibility into Electric Generation Planning. Impacts and Methods for System Design and Policy Analysis*. PhD thesis, Engineering Systems Division. Massachusetts Institute of Technology.

Palmintier, B. and Webster, M. (2011). Impact of unit commitment constraints on generation expansion planning with renewables. Detroit, MI. IEEE Power and Energy Society.

Palmintier, B. and Webster, M. (2013). Impact of operational flexibility on generation planning.

Papavasiliou, A., Oren, S. S., and O'Neill, R. P. (2011). Reserve requirements for wind power integration: A scenario-based stochastic programming framework. *IEEE Transactions on Power Systems*, 26(4):2197–2206.

Parsons, J. E. and Mello, A. S. (2012). *Lecture Notes on Advance Corporate Financial Risk*. MIT.

Pereira, M. V. F. and Pinto, L. M. V. G. (1985). Stochastic optimization of a multireservoir hydroelectric system: A decomposition approach. *Water Resource Research*, 21(6):779–792.

Perez-Arriaga, I. J. (1994). *Principios Economicos Marginalistas en los Sistemas de Energia Electrica*. Madrid, Spain.

Perez-Arriaga, I. J. (2007). Security of electricity supply in europe in a short, medium and long-term perspective. *European Journal of Energy Markets*, 2(2).

Perez-Arriaga, I. J. (2011). Managing large scale penetration of intermittent renewables. Mitei symposium (p. 43), MIT, Cambridge, MA.

Perez-Arriaga, I. J. and Batlle, C. (2012). Impacts of intermittent renewables on electricity generation system operation. *Econmics of Energy and Environmental Policy*, 1(3):17.

Perez-Arriaga, I. J. and Linares, P. (2008). Markets vs. regulation: A role for indicative energy planning. *Energy Journal*, page 13.

Perez-Arriaga, I. J. and Meseguer, C. (1997). Wholesale marginal prices in competitve generation markets. *IEEE Transactions on Power Systems*, 12(2).

Poyry and EWEA (2010). *Wind Energy and Electricity Prices. Exploring the 'Merit Order' Effect*. European Wind Energy Association.

Rodilla, P. (2010). Regulatory tools to enhance security of supply at the generation level in electricity markts.

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

Rodilla, P., Cerisola, S., and Batlle, C. (2013). Modeling the major overhaul cost of gas-fired plants in the unit commitment problem. *IEEE Transactions on Power Systems*, PP(99):1–11.

Rothwell, G. and Gomez, T. (2003). *Electricity Economics. Regulation and Deregulation*. Wiley Inter-Science, Piscataway, NJ.

Rubinstein, R. Y. and Kroese, D. P. (2008). *Simulation and the Monte Carlo Method*. John Wiley and Sons, Hoboken, NJ.

Schweppe, F. C., Caramanis, M. C., Tabors, R. D., and Bohn, R. E. (1987). *Spot Pricing of Electricity*. Kluwer Academic Publishers, Boston.

Sioshansi, F. P. (2006). *Electricity Market Reform: An International Perspective*. Elsevier.

Sioshansi, F. P. (2008). *Competitive Electricity Markets. Design, Implementation, Performance*. Elsevier, CA.

Stoft, S. (2002). *Power System Economics: Designing Markets for Electricity*. Wiley Inter-Science.

Takriti, S., Birge, J. R., and Long, E. (1996). A stochastic model for the unit commitment problem. *IEEE Transactions on Power Systems*, 11(3):1497–1508.

Troy, N., Denny, E., and O'Malley, M. (2010). Base-load cycling on a system with significant wind penetration. *IEEE Transactions on Power Systems*, 25(2):1088–1097.

Tuohy, A., Meibom, P., Denny, E., and O'Malley, M. (2009). Unit commitment for systems with significant wind penetration. *IEEE Transactions on Power Systems*, 24(2):592–601.

Turvey, R. and Anderson, D. (1977). *Electricity Economics. Essays and Case Studies*. The World Bank, The Johns Hopkins University Press, Baltimore and London.

Vazquez, C. (2003). *Modelos de Casacion de Ofertas en Mercados Electricos*. PhD thesis, Universdad Pontificia de Comillas - IIT.

Ventosa, M., Baillo, A., Ramos, A., and Rivier, M. (2005). Electricity market modeling trends. *Energy Policy*, 33:897–913.

Appendix A

Parameters Used in IMRES

Tables A.1, A.2 & A.3 present the parameters used for the thermal technologies included in the analyses:

Table A.1: Fixed Costs of Thermal Power Plants

Tech.	Capital Cost [k$/MW]	Life [years]	WACC [%]	Fixed O&M [$/MW-year]	Annualized Capital Cost [k$/MW year]
Nuclear	5,335	40	7	88,750	489
Coal	3,167	40	10	35,970	360
CCGT	978	20	10	14,390	129
OCGT	974	40	10	6,980	106

Table A.2: Variable Costs of Thermal Power Plants

Tech.	Variable O&M [$/MWh]	Heat rate [MBTU/MWh]	Fuel price [$/MBTU]	Variable cost [$/MWh]
Nuclear	2.04	10.49	0.43	6.5
Coal	4.25	8.8	2.22	23.8
CCGT	3.43	7.05	7.81	58.5
OCGT	14.7	10.85	7.81	99.4

Table A.3: Technical Parameters of Thermal Power Plants

Tech.	P_i^{MAX} [GW]	P_i^{MIN} [GW]	R_i [GW/h]	C_i^{STUP} [M$/stup]	M_i^U [hrs]	M_i^D [hrs]
Nuclear	1	0.9	0.19	1	36	36
Coal	0.5	0.35	0.21	0.125	6	6
CCGT	0.4	0.15	0.32	0.06	3	3
OCGT	0.2	0.05	0.36	0.01	0	0

Appendix B

Multi-Period Discounting

The reader might be familiar already with the single-year version of discounting:

$$PV_S = \sum_{t=0}^{T} \frac{CF_t}{(1 + r_S)^t} \tag{B.1}$$

If payments are evenly spread throughout the year, having a nominal rate r_C compounded m times per year, and we assume that cash flows are received at the beginning of each year, we have that:

$$PV_C = \sum_{t=0}^{T} \frac{CF_t}{(1 + \frac{r_C}{m})^{mt}} \tag{B.2}$$

If we make the compounding period infinitesimally small, using the definition of the exponential number $e^r = \lim m \to \infty (1 + r/m)^m$, it follows that:

$$PV_I = \sum_{t=0}^{T} \frac{CF_t}{\lim_{m\to\infty}(1 + r_I/m)^{mt}} = \sum_{t=0}^{T} CF_t e^{-r_I t} \tag{B.3}$$

From the previous expressions, it also follows the equivalence between the nominal rate used in continuous compounding, r_I, and the single year rate of return, r_S, is: $e^{r_I} = (1 + r_S)$. The method described in this paper uses the generalized case of continuous compounding, and assume that a year is divided into n periods, that cash flows are received at each of

those n periods, and that compounding occurs continuously during the period:

$$PV = \sum_{i=0}^{N} CF_i e^{-ri/n}, \qquad \text{with} \quad N = nT \tag{B.4}$$

Appendix C

PowerCo's Example

Figure C-1: PowerCo: Valuation Using a Risk-Adjusted Discount Rate (Parsons and Mello, 2012).

Method #1: Risk Adjusted Discount Rate Method -- simultaneously adjust for risk and time						
[1]	Year	1	2	3	4	5
[2]	Production (000 units)	5,400	5,700	6,200	7,000	8,100
[3]	Unit Price	1.21	1.47	1.78	2.15	2.60
[4]	Revenue ($ 000)	6,545	8,369	11,021	15,058	21,073
[5]	Risk-adjusted Discount Rate, r_a	11%	11%	11%	11%	11%
[6]	Risk-adjusted Discount Factor	0.896	0.803	0.720	0.646	0.579
[7]	PV ($ 000)	5,864	6,720	7,935	9,724	12,211
[8]	Total PV ($ 000)	42,455				
Method #2: Certainty Equivalent Method -- separately adjust for risk then for time						
[9]	Revenue ($ 000)	6,545	8,369	11,021	15,058	21,073
[10]	Certainty Equivalent Risk Premium, λ	6%	6%	6%	6%	6%
[11]	Certainty Equivalent Discount Factor	0.942	0.887	0.836	0.789	0.744
[12]	Certainty Equivalent Revenue ($ 000)	6,165	7,427	9,218	11,873	15,670
[13]	Risk-free Discount Rate, r_f	5.0%	5.0%	5.0%	5.0%	5.0%
[14]	Risk-free Discount Factor	0.951	0.905	0.861	0.819	0.779
[15]	PV ($ 000)	5,864	6,720	7,935	9,724	12,211
[16]	Total PV ($ 000)	42,455				

Notes:
[2]=given.
[3]=given.
[4]=[2]*[3].
[5]=given.
[6]=exp(-[5]*[1])
[7]=[4]*[6]
[8]=sum([7])
[9]=[4].
[10]=[5]-[13].
[11]=exp(-[10]*[1])
[12]=[9]*[11]
[13]=given.
[14]=exp(-[13]*[1]).
[15]=[12]*[14].
[16]=sum([15])

Figure C-2: PowerCo: The Macro Perspective – the Underlying Risk Factor Model (Parsons and Mello, 2012).

[1]	Year	1	2	3	4	5
[2]	Cum Market Return, annualized					
	0	47%	35%	30%	27%	24%
	1	31%	23%	20%	18%	17%
	2	14%	12%	11%	10%	10%
	3	-2%	1%	2%	2%	3%
	4	-18%	-11%	-8%	-6%	-4%
	5	-34%	-22%	-17%	-14%	-12%
[3]	Market Index (Initial Value = 1)					
	0	1.594	2.008	2.432	2.888	3.385
	1	1.357	1.599	1.840	2.093	2.362
	2	1.155	1.273	1.393	1.517	1.648
	3	0.984	1.014	1.054	1.099	1.150
	4	0.837	0.808	0.797	0.797	0.802
	5	0.713	0.643	0.603	0.577	0.560
[4]	Probabilities					
	0	3.1%	3.1%	3.1%	3.1%	3.1%
	1	15.6%	15.6%	15.6%	15.6%	15.6%
	2	31.3%	31.3%	31.3%	31.3%	31.3%
	3	31.3%	31.3%	31.3%	31.3%	31.3%
	4	15.6%	15.6%	15.6%	15.6%	15.6%
	5	3.1%	3.1%	3.1%	3.1%	3.1%
[5]	Expected Market Index	1.083	1.174	1.271	1.377	1.492
[6]	Expected Market Return, annualized	6.4%	6.4%	6.4%	6.4%	6.4%
[7]	Volatility of Market Return, annualized	18.0%	18.0%	18.0%	18.0%	18.0%
[8]	State-Contingent Risk Premia, annualized					
	0	38.7%	27.8%	23.0%	20.2%	18.2%
	1	23.8%	17.3%	14.4%	12.7%	11.5%
	2	8.9%	6.7%	5.7%	5.2%	4.8%
	3	-6.1%	-3.9%	-2.9%	-2.4%	-2.0%
	4	-21.0%	-14.5%	-11.6%	-9.9%	-8.7%
	5	-35.9%	-25.1%	-20.2%	-17.4%	-15.4%
[9]	State-Contingent Discount Factors					
	0	0.679	0.573	0.501	0.446	0.402
	1	0.788	0.708	0.650	0.602	0.563
	2	0.915	0.875	0.842	0.814	0.788
	3	1.063	1.081	1.092	1.099	1.103
	4	1.234	1.336	1.415	1.483	1.544
	5	1.433	1.650	1.835	2.003	2.161
[10]	Certainty Equivalent Value of Market Index	1.05	1.11	1.16	1.22	1.28
[11]	Risk-free Discount Rate, r_f	5.0%	5.0%	5.0%	5.0%	5.0%
[12]	Risk-free Discount Factor	0.951	0.905	0.861	0.819	0.779
[13]	PV of Market Index	1.00	1.00	1.00	1.00	1.00

Figure C-3: PowerCo: The Micro Perspective – Asset Exposure (Parsons and Mello, 2012).

[1]	Year	1	2	3	4	5
[2]	Production (000 units)	5,400	5,700	6,200	7,000	8,100
[3]	Unit Price					
	0	2.54	4.03	5.92	8.34	11.46
	1	1.84	2.56	3.39	4.38	5.58
	2	1.33	1.62	1.94	2.30	2.72
	3	0.97	1.03	1.11	1.21	1.32
	4	0.70	0.65	0.64	0.63	0.64
	5	0.51	0.41	0.36	0.33	0.31
[4]	Revenue at PowerCo, ($ 000)					
	0	13,724	22,976	36,678	58,382	92,831
	1	9,946	14,572	20,999	30,662	45,186
	2	7,208	9,242	12,022	16,103	21,994
	3	5,224	5,861	6,883	8,457	10,706
	4	3,786	3,717	3,941	4,442	5,211
	5	2,743	2,358	2,256	2,333	2,536
[5]	Probabilities					
	0	3.1%	3.1%	3.1%	3.1%	3.1%
	1	15.6%	15.6%	15.6%	15.6%	15.6%
	2	31.3%	31.3%	31.3%	31.3%	31.3%
	3	31.3%	31.3%	31.3%	31.3%	31.3%
	4	15.6%	15.6%	15.6%	15.6%	15.6%
	5	3.1%	3.1%	3.1%	3.1%	3.1%
[6]	Expected Unit Price	1.21	1.47	1.78	2.15	2.60
[7]	Expected Revenue ($ 000)	6,545	8,369	11,021	15,058	21,073

Figure C-4: PowerCo: Valuation by the State-Contingent Discounting Method (Parsons and Mello, 2012).

[1]	Year	1	2	3	4	5
[2]	State Contingent Revenue ($ 000)					
	0	13,724	22,976	36,678	58,382	92,831
	1	9,946	14,572	20,999	30,662	45,186
	2	7,208	9,242	12,022	16,103	21,994
	3	5,224	5,861	6,883	8,457	10,706
	4	3,786	3,717	3,941	4,442	5,211
	5	2,743	2,358	2,256	2,333	2,536
[3]	Probabilities					
	0	3.1%	3.1%	3.1%	3.1%	3.1%
	1	15.6%	15.6%	15.6%	15.6%	15.6%
	2	31.3%	31.3%	31.3%	31.3%	31.3%
	3	31.3%	31.3%	31.3%	31.3%	31.3%
	4	15.6%	15.6%	15.6%	15.6%	15.6%
	5	3.1%	3.1%	3.1%	3.1%	3.1%
[4]	State-Contingent Discount Factors					
	0	0.679	0.573	0.501	0.446	0.402
	1	0.788	0.708	0.650	0.602	0.563
	2	0.915	0.875	0.842	0.814	0.788
	3	1.063	1.081	1.092	1.099	1.103
	4	1.234	1.336	1.415	1.483	1.544
	5	1.433	1.650	1.835	2.003	2.161
[5]	Certainty Equivalent Revenue ($ 000)	6,165	7,427	9,218	11,873	15,670
[6]	Risk-free Discount Rate, r_f	5.0%	5.0%	5.0%	5.0%	5.0%
[7]	Risk-free Discount Factor	0.951	0.905	0.861	0.819	0.779
[8]	PV ($ 000)	5,864	6,720	7,935	9,724	12,211
[9]	Total PV ($ 000)	42,455				

Figure C-5: PowerCo: Valuing Partnership #1(Parsons and Mello, 2012).

[1]	Year	1	2	3	4	5
[2]	Partnership #1 Unit Price Target	0.37	0.37	0.37	0.37	0.37
[3]	Partnership #1 Revenue ($ 000)					
	0	1,998	2,109	2,294	2,590	2,997
	1	1,998	2,109	2,294	2,590	2,997
	2	1,998	2,109	2,294	2,590	2,997
	3	1,998	2,109	2,294	2,590	2,997
	4	1,998	2,109	2,294	2,590	2,997
	5	1,998	2,109	2,256	2,333	2,536
[4]	Probabilities					
	0	3.1%	3.1%	3.1%	3.1%	3.1%
	1	15.6%	15.6%	15.6%	15.6%	15.6%
	2	31.3%	31.3%	31.3%	31.3%	31.3%
	3	31.3%	31.3%	31.3%	31.3%	31.3%
	4	15.6%	15.6%	15.6%	15.6%	15.6%
	5	3.1%	3.1%	3.1%	3.1%	3.1%
[5]	State-Contingent Discount Factors					
	0	0.679	0.573	0.501	0.446	0.402
	1	0.788	0.708	0.650	0.602	0.563
	2	0.915	0.875	0.842	0.814	0.788
	3	1.063	1.081	1.092	1.099	1.103
	4	1.234	1.336	1.415	1.483	1.544
	5	1.433	1.650	1.835	2.003	2.161
[6]	Certainty Equivalent Net Cash Flow ($ 000)	1,998	2,109	2,292	2,574	2,966
[7]	Risk-free Discount Rate, r_f	5.0%	5.0%	5.0%	5.0%	5.0%
[8]	Risk-free Discount Factor	0.951	0.905	0.861	0.819	0.779
[9]	PV ($ 000)	1,900	1,908	1,973	2,108	2,311
[10]	Total PV ($ 000)	10,201				

Figure C-6: PowerCo: Valuing Partnership #2 (Parsons and Mello, 2012).

[1]	Year	1	2	3	4	5
[2]	Partnership #1 Unit Price Target	0.37	0.37	0.37	0.37	0.37
[3]	Partnership #2 Revenue ($ 000)					
	0	11,726	20,867	34,384	55,792	89,834
	1	7,948	12,463	18,705	28,072	42,189
	2	5,210	7,133	9,728	13,513	18,997
	3	3,226	3,752	4,589	5,867	7,709
	4	1,788	1,608	1,647	1,852	2,214
	5	745	249	0	0	0
[4]	Probabilities					
	0	3.1%	3.1%	3.1%	3.1%	3.1%
	1	15.6%	15.6%	15.6%	15.6%	15.6%
	2	31.3%	31.3%	31.3%	31.3%	31.3%
	3	31.3%	31.3%	31.3%	31.3%	31.3%
	4	15.6%	15.6%	15.6%	15.6%	15.6%
	5	3.1%	3.1%	3.1%	3.1%	3.1%
[5]	State-Contingent Discount Factors					
	0	0.679	0.573	0.501	0.446	0.402
	1	0.788	0.708	0.650	0.602	0.563
	2	0.915	0.875	0.842	0.814	0.788
	3	1.063	1.081	1.092	1.099	1.103
	4	1.234	1.336	1.415	1.483	1.544
	5	1.433	1.650	1.835	2.003	2.161
[6]	Certainty Equivalent Net Cash Flow ($ 000)	4,167	5,318	6,927	9,299	12,704
[7]	Risk-free Discount Rate, r_f	5.0%	5.0%	5.0%	5.0%	5.0%
[8]	Risk-free Discount Factor	0.951	0.905	0.861	0.819	0.779
[9]	PV ($ 000)	3,963	4,812	5,962	7,616	9,900
[10]	Total PV ($ 000)	32,254				

Appendix D

The Black-Scholes Model as the Limit of the Binomial Model

In this appendix we prove using continuously-compounded discounting that the Black-Scholes model is the limit of a binomial distribution with probabilities π and $1-\pi$, and returns equal to u and d respectively. This proof is based on the proof found in (Ingersoll, 1987), but expressed in terms of the continuously-compounded discounting notation, instead of discretely-compounded. The same proof with a different notation can be found in other textbooks such as (Cox and Rubinstein, 1985).

Let's assume that the continuously-compounded rate of return on some stock S follows a continuous-time random walk, $dR = (\mu - \frac{\sigma^2}{2})dt + \sigma dW$, where dW is the normally distributed incremental step in a Brownian motion. The equivalent discrete-time distribution of the random walk is $\Delta R_i = \log(S_i/S_{i-1}) = (\mu - \frac{\sigma^2}{2})\Delta t + \sigma\varepsilon\sqrt{\Delta t}$, such that the returns on the stock are distributed log-normally, $S_i/S_{i-1} \sim \ln\mathcal{N}((\mu - \frac{\sigma^2}{2})\Delta t, \sigma\sqrt{\Delta t})$, with an expected return over one period [1] $\mathbb{E}[S_i/S_{i-1}] = e^{\mu\Delta t}$, and variance $\mathrm{Var}[S_i/S_{i-1}] = e^{2\mu\Delta t}(e^{\sigma^2\Delta t} - 1)$.

[1] Note that if a random variable $X \sim \mathcal{N}(\mu, \sigma)$, then $e^X \sim \ln\mathcal{N}(\mu, \sigma)$, and $\mathbb{E}[e^X] = e^{\mu+\sigma^2/2}$. Some options textbooks like (Ingersoll, 1987) and (Cox and Rubinstein, 1985) use discretely-compounded discounting, and the returns on the stock and the rate of return follow a similar probability distribution because $S/S_0 = 1+R$. Here, as we use continuously-compounded discounting, with returns log-normally distributed, the underlying normal probability has to be shifted down by $\sigma^2/2$ so that the expected returns of the log-normal distribution have a mean of $e^{\mu\Delta t}$. Yet, there are two equivalent ways to write the same continuous process derived here. In one, μ is defined as the variable that is the continuous rate of drift, i.e., the parameter multiplying Δt. In

Equations (D.1-D.3) show how the two moments of the process are derived:

$$\mathbb{E}[S_i/S_{i-1}] = \mathbb{E}[e^{(\mu-\frac{\sigma^2}{2})\Delta t+\sigma\varepsilon\sqrt{\Delta t}}] = e^{(\mu-\frac{\sigma^2}{2})\Delta t}\mathbb{E}[e^{\sigma\varepsilon\sqrt{\Delta t}}] = e^{\mu\Delta t} \tag{D.1}$$

$$\mathbb{E}[(S_i/S_{i-1})^2] = \mathbb{E}[e^{2(\mu-\frac{\sigma^2}{2})\Delta t+2\sigma\varepsilon\sqrt{\Delta t}}] = e^{(2\mu-\sigma^2)\Delta t}\mathbb{E}[e^{2\sigma\varepsilon\sqrt{\Delta t}}] = e^{(2\mu+\sigma^2)\Delta t} \tag{D.2}$$

$$\mathrm{Var}[S_i/S_{i-1}] = \mathbb{E}[(S_i/S_{i-1})^2] - \mathbb{E}[S_i/S_{i-1}]^2 = e^{(2\mu+\sigma^2)\Delta t} - e^{2\mu\Delta t} = e^{2\mu\Delta t}(e^{\sigma^2\Delta t} - 1) \tag{D.3}$$

We now define a binomial distribution with the same mean and variance of the log-normal process just described. Let π be the binomial probability of a tree, where u and d are the returns of each of the two branches of the tree, such that $S_u = uS_{i-1}$ and $S_d = dS_{i-1}$. We now impose on the tree the condition that its two moments are the same as those of the log-normal process: $\mathbb{E}[S_i/S_{i-1}] = e^{\mu\Delta t}$ and $\mathrm{Var}[S_i/S_{i-1}] = e^{2\mu\Delta t}(e^{\sigma^2\Delta t} - 1)$.

$$\begin{cases} \mathbb{E}[S_i/S_{i-1}] = u\pi + d(1-\pi) = e^{\mu\Delta t} & \text{(D.4a)} \\ \mathrm{Var}[S_i/S_{i-1}] = (u-d)^2\pi(1-\pi) = e^{2\mu\Delta t}(e^{\sigma^2\Delta t} - 1) & \text{(D.4b)} \end{cases}$$

Solving the system of equations for the values of u and d, we have that:

$$u = e^{\mu\Delta t}\left[1 + (1-\pi)\left(\frac{e^{\sigma^2\Delta t} - 1}{\pi(1-\pi)}\right)^{\frac{1}{2}}\right] \tag{D.5}$$

$$d = e^{\mu\Delta t}\left[1 - \pi\left(\frac{e^{\sigma^2\Delta t} - 1}{\pi(1-\pi)}\right)^{\frac{1}{2}}\right] \tag{D.6}$$

This binomial distribution, in the limit, approximates the continuous log-normal distribution $\ln\mathcal{N}((\mu - \frac{\sigma^2}{2})\Delta t, \sigma\sqrt{\Delta t})$.

We now define the risk-neutral distribution of the binomial tree, such that the expected return on the stock with respect to the risk neutral probability is the risk free rate of return: $\pi^* = \frac{e^{r\Delta t}-d}{u-d}$. A risk neutral investor would be indifferent between an investment with a certain rate of return equal to $r\Delta t$ and the uncertain return in the risk neutral tree, yielding

that formulation, the annual expected stock value is given using $\mu + \frac{\sigma^2}{2}$. In the second formulation, μ is the return on the expected stock price. Then the parameter multiplying Δt is $\mu - \frac{\sigma^2}{2}$. The original formulation back in the time of Black-Scholes was the latter (the one used in this appendix), but both are equivalent.

the same expected returns. The value of π^* can be determined substituting the values of u and d by their respective expressions derived above:

$$\pi^* = \frac{e^{r\Delta t} - d}{u - d} = \frac{e^{r\Delta t} - e^{\mu\Delta t}\left[1 - \pi\left(\frac{e^{\sigma^2\Delta t}-1}{\pi(1-\pi)}\right)^{\frac{1}{2}}\right]}{e^{\mu\Delta t}\left(\frac{e^{\sigma^2\Delta t}-1}{\pi(1-\pi)}\right)^{\frac{1}{2}}} = \pi + \left(\frac{e^{r\Delta t}}{e^{\mu\Delta t}} - 1\right)\left(\frac{e^{\sigma^2\Delta t}-1}{\pi(1-\pi)}\right)^{-\frac{1}{2}} \quad \text{(D.7)}$$

We now solve for the mean and variance of the binomial risk-neutral distribution.

$$\begin{aligned}
\mathbb{E}^*[S_i/S_{i-1}] \quad &= u\pi^* + d(1-\pi^*) = \pi^*(u-d) + d = \\
&= e^{r\Delta t} - e^{\mu\Delta t} + \pi e^{\mu\Delta t}\left(\frac{e^{\sigma^2\Delta t}-1}{\pi(1-\pi)}\right)^{\frac{1}{2}} + e^{\mu\Delta t} - \pi e^{\mu\Delta t}\left(\frac{e^{\sigma^2\Delta t}-1}{\pi(1-\pi)}\right)^{\frac{1}{2}} \quad \text{(D.8)} \\
&= e^{r\Delta t}
\end{aligned}$$

$$\begin{aligned}
\mathrm{Var}^*[S_i/S_{i-1}] \quad &= \quad u^2\pi^* + d^2(1-\pi^*) - e^{2r\Delta t} = d^2 + \pi^*(u^2-d^2) - e^{2r\Delta t} \quad \text{(D.9)} \\
&= \quad e^{2\mu\Delta t}\left[1 + \pi\frac{e^{\sigma^2\Delta t}-1}{1-\pi} - 2\pi\left(\frac{e^{\sigma^2\Delta t}-1}{\pi(1-\pi)}\right)^{\frac{1}{2}}\right] + \\
&\quad + \left[\pi + \left(\frac{e^{r\Delta t}}{e^{\mu\Delta t}} - 1\right)\left(\frac{e^{\sigma^2\Delta t}-1}{\pi(1-\pi)}\right)^{-\frac{1}{2}}\right] \cdot \\
&\quad \cdot e^{2\mu\Delta t}\left[2\left(\frac{e^{\sigma^2\Delta t}-1}{\pi(1-\pi)}\right)^{\frac{1}{2}} + (1-2\pi)\frac{e^{\sigma^2\Delta t}-1}{\pi(1-\pi)}\right] - e^{2r\Delta t} = \\
&= \quad e^{2\mu\Delta t} + e^{2\mu\Delta t}\pi\frac{e^{\sigma^2\Delta t}-1}{1-\pi} - 2\pi e^{2\mu\Delta t}\left(\frac{e^{\sigma^2\Delta t}-1}{\pi(1-\pi)}\right)^{\frac{1}{2}} + \\
&\quad + 2e^{\mu\Delta t}e^{r\Delta t} + (1-2\pi)e^{\mu\Delta t}e^{r\Delta t}\left(\frac{e^{\sigma^2\Delta t}-1}{\pi(1-\pi)}\right)^{\frac{1}{2}} - \\
&\quad - 2e^{2\mu\Delta t} - (1-2\pi)e^{2\mu\Delta t}\left(\frac{e^{\sigma^2\Delta t}-1}{\pi(1-\pi)}\right)^{\frac{1}{2}} +
\end{aligned}$$

$$
\begin{aligned}
&\quad + 2\pi e^{2\mu\Delta t}\left(\frac{e^{\sigma^2\Delta t}-1}{\pi(1-\pi)}\right)^{\frac{1}{2}} + e^{2\mu\Delta t}\pi(1-2\pi)\frac{e^{\sigma^2\Delta t}-1}{\pi(1-\pi)} - e^{2r\Delta t} = \\
= &\quad - e^{2\mu\Delta t} + e^{2\mu\Delta t}(e^{\sigma^2\Delta t}-1) + 2e^{\mu\Delta t}e^{r\Delta t} + \\
&\quad + (1-2\pi)e^{\mu\Delta t}\left(\frac{e^{\sigma^2\Delta t}-1}{\pi(1-\pi)}\right)^{\frac{1}{2}}\left(e^{r\Delta t}-e^{\mu\Delta t}\right) - e^{2r\Delta t} = \\
(\text{if } \pi = 0.5) \quad = &\quad e^{2\mu\Delta t}(e^{\sigma^2\Delta t}-1) - \left(e^{r\Delta t}-e^{\mu\Delta t}\right)^2
\end{aligned}
$$

In the limit, this risk-neutral binomial process defines the continuous risk-neutral distribution, which inherits the mean and variance of the binomial process. We know that the variance of a log-normal distribution with $\mathbb{E}^*[S_i/S_{i-1}] = e^{r\Delta t}$ is $\text{Var}^*[S_i/S_{i-1}] = e^{2r\Delta t}(e^{\sigma^{*2}\Delta t}-1)$, where σ^* is the standard deviation of the lognormal distribution. If we equate this variance to the variance of the risk-neutral binomial process we have that:

$$
\begin{aligned}
\sigma^* &= \left[\ln\left(\frac{e^{2\mu\Delta t}}{e^{2r\Delta t}}\left(e^{\sigma^2\Delta t}-1\right) - \frac{1}{e^{2r\Delta t}}\left(e^{r\Delta t}-e^{\mu\Delta t}\right)^2 + 1\right)\right]^{\frac{1}{2}} = \qquad \text{(D.10)} \\
&= \left[\ln\left(\frac{e^{2\mu\Delta t}e^{\sigma^2\Delta t} - 2e^{2\mu\Delta t} + 2e^{r\Delta t}e^{\mu\Delta t}}{e^{2r\Delta t}}\right)\right]^{\frac{1}{2}} = \\
(\text{if } e^{(\mu-r)} \sim 1) \quad &\sim \sigma
\end{aligned}
$$

Therefore, the underlying lognormal risk-neutral process is $\ln\mathcal{N}((r-\frac{\sigma^2}{2})\Delta t, \sigma\sqrt{\Delta t})$. $\square$