JOHNS HOPKINS UNIVERSITY

OPTIMIZATION MODELING TO ADDRESS THE IMPACTS OF ELECTRIC POWER MARKET DESIGN ON OPERATIONS AND POLICY

by

Robin Sarah Broder Hytowitz

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland

April 2018

ABSTRACT

The electric power system is affected by numerous inefficiencies. Operation of the power grid uses intricate mathematical models to schedule supply and demand instantaneously, and complex settlement mechanisms to charge and pay participants. This dissertation focuses on four aspects of electric market design and operation endeavoring to improve economic and operational efficiency. Each chapter utilizes bottom-up engineering-economic models to simulate power grid operations. The overall goal of the dissertation is to analyze electric market inefficiencies and examine proposed alternative designs and policies.

The dissertation begins with characterizing the electric system and the role and challenges of renewable energy in Chapter 1. Then Chapter 2 proposes a new method for pricing electricity in organized wholesale markets, called the Dual Pricing Algorithm. The current pricing method is non-confiscatory but does not capture the full cost of operation in marginal prices. The proposed method achieves these two aims while also providing further transparency. Chapter 3 examines potential benefits of three adjustments in reserve procurement procedures, and estimates economic efficiency using a European test system. Each adjustment improves current practice, either in the quantity of reserves needed, the procurement method, or the degree of coordination with neighboring countries. The results demonstrate coordination among countries shows greatest consistent benefits among the three adjustments. Chapter 4 examines integration of carbon policies into real-time markets when the emissions system encompasses a sub-region of the larger electricity market, comparing five alternative models. Findings suggest that there is a trade-off between emissions and cost, with no one dominant method to identify and manage leakage from the regulated system. Chapter 5 analyzes degrees of coordination between

neighboring systems for both day-ahead and real-time energy markets. The simulations for a test case find that coordinating in real-time without coordination in the day-ahead market results in higher costs compared to not coordinating at all.

These chapters examine trade-offs, whether they are between ease of implementation, economic efficiency, renewable integration, or emissions reductions. Overall, the dissertation contributes a framework for assessing market design improvements, and demonstrates to system operators and decisions makers that coordination between neighboring regions can increase economic efficiency.

Academic Advisor and Reader:

Benjamin F. Hobbs, Schad Professor of Environmental Management
Department of Environmental Health and Engineering
The Johns Hopkins University

Reader:

Sauleh Siddiqui, Assistant Professor
Department of Civil Engineering
The Johns Hopkins University

Reader:

John Boland, Professor Emeritus
Department of Environmental Health and Engineering
The Johns Hopkins University

DEDICATION

For my husband, Jon: it has been a long road that took us around the world, thank you for being up for the adventure.

ACKNOWLEDGEMENTS

This dissertation has been a long journey, and I am eternally grateful to my academic advisers, my family, friends near and far, and coffee for helping me along the way. I would not have been able to write this dissertation without my academic adviser, Ben Hobbs. His advice has been invaluable through the PhD process, and I am grateful for the many hours of feedback and guidance through the vastly complex energy systems field. I do not think I ever found a topic where he was not knowledgeable, with at least four published papers. Thank you for pushing me to do great work and having eternal optimism.

I would also like to thank my readers, John Boland and Sauleh Siddiqui. Prof. Boland pushed me to think about the economic impact of my work during my department qualifying exam, which helped me through this dissertation. Prof. Siddiqui has taught me a great deal throughout my PhD, beginning in a GAMS workshop in my first year, through his class on multilevel optimization, and in my qualifying exam. I am grateful to Profs. Boland and Siddiqui for their time and efforts.

Chapter 2 is a collaboration with Richard O'Neill and staff at the Federal Energy Regulatory Commission. My PhD experience would not have been the same without the guidance and tutelage of Dick, I have learned a great deal about the history of the industry and the regulatory side of electricity. The Office of Energy Policy and Innovation welcomed me with open arms, and I am thankful I was able to intern with them. I am also grateful for my many FERC colleagues who have let me pick their brains over the years about market issues: Emma, Dan, Elisabeth, Pete, Dave, Jon, Rodger, and Samin.

Chapter 3 is a collaboration with Özge Özdemir and the Energy Research Center of the Netherlands (ECN), and funded by the NSF WINDINSPIRE grant. The grant enabled me to travel to the Netherlands several times throughout my PhD, and I am grateful to have had the international experience. I would also like to thank my ECN colleagues Marit and Paul for their comments and assistance over the years. I am especially indebted to Özge, who has worked with me beginning in my first year. We spent many hours struggling through code and I have learned an immense amount from her. I am also thankful to Kenneth Van den Bergh for the joint collaboration and enlightening discussions about European reserves.

Chapter 4 is a collaboration with Prof. Harry van der Weijde at the University of Edinburgh in Scotland. He was greatly helpful during the formulation and modeling. Research for Chapter 5 began during a Fulbright Fellowship to Denmark at the Centre for Electric Power, Technical University of Denmark. I am very appreciative of my mentors, Prof. Pierre Pinson and Prof. Qiuwei Wu, and colleagues from DTU: Tue, Jakob, Pauli, Daniel, Johan, and Stefanos. Chapter 4 and Chapter 5 are also funded by NSF WINDINSPIRE.

I am also extremely grateful to my fellow labmates (Andrea, Anya, Brent, Carl, Cindy, Elina, Emily, Fengwei, Haralampos, Liang, Marc, Pearl, Pengcheng, Qingyu, Roderick, Saamrat, and Venkat) and classmates (Dano, Mike, Shane, Steven, and many more) in DoGEE. I am very lucky to have Anya Castillo and Brent Eldridge as fellow students and fellow FERC employees, their support helped me in balancing both jobs. Elina Spyrou has been a remarkable classmate, officemate, brainstormer, and researcher. I cannot thank her enough for the past four years of friendship and encouragement, her words of wisdom helped me immeasurably through times of frustration and stress. I look forward to future trips, celebrations, and even collaborations. I am also thankful to Prof. Kory Hedman and his group at Arizona State University for sparking my interest in markets during my master's degree. Nikita and Mojdeh have both been inspirations to me years after leaving ASU, and I am looking forward to the next chapter.

Over the past several years I have been lucky to find inquisitive, creative, and witty friends in Baltimore who also love playing trivia. I am very thankful for Jazz Fingers: Deb, Rick, Andrea, and Jay (and now Leah). From Shenandoah to Havre de Grace, I have loved exploring, eating and drinking across the mid-Atlantic. My time in Baltimore was made better because of the hours of entertaining chats, local scavenger hunts, Star Wars watch parties, and many celebrations in between.

Finally, I would like to thank my family and friends. Over the years my uncles, aunts, and siblings have been more than supportive, and I am looking forward to spending much more time with them. None of this would have been possible without my parents, Susan and Mel. Throughout my life they have supported and encouraged me to accomplish any goal I set. My father is always willing to talk through a complex topic, starting when I asked him to define energy in elementary school. My mother has been my constant inspiration; she taught me to stand up for myself and pushes me still to find my voice. Whenever I needed encouragement throughout my many, many years of school, she was on the phone telling me to 'kick some butt.' My gratitude to them is everlasting.

Lastly, I am not sure if I could have accomplished all that I have my partner and best friend, Jon. He moved with me to New York, Phoenix, Copenhagen, Baltimore, and now West again. Among many things, thank you for making me chili and doing laundry when I was busy with work. I am so very grateful and lucky to have such an amazing partner.

This work is funded by NSF PIRE Grant OISE 1243482 and the Federal Energy Regulatory Commission.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

NOMENCLATURE

Sets

D	All demand or customers in the system
D^+	Demand dispatched in the unit commitment solution, subset of *D*
D^0	Demand not dispatched in the unit commitment solution, subset of *D*
G	All generators in the system (*g*), where G_i indicates the set of generators at bus *i*
G^{fast}	Fast generators, subset of *G*
G^{slow}	Slow generators, subset of *G*
G^+	Generators dispatched in the unit commitment solution, subset of *G*
G^{NL}	Generators in the Netherlands, subset of *G*
G^{EU}	Generators in Europe except the Netherlands, subset of *G*
N	All buses in the system (*i*)
K	All lines in the system (*k*)
K^{IT}	Intertie lines connecting systems *A* and *B*
T	Time (*t*)
V	All types of storage technologies in the system, where V_i indicates the set of storage at bus *i* (v)
$\delta^{+/-}(i)$	Set of lines defined as connected to (+) or from (-) bus *i*
$i(k), j(k)$	Set of originating (*i*) and ending (*j*) nodes that composes line *k*

Variables

d_{it}	Dispatch for demand at bus *i* in period *t*
$d_{i,t}^{\text{curt}}$	Demand curtailed a bus *i* in period *t*
$p_{g,t}$	Power output from generator *g* in period *t*
$f_{k,t}$	Power flowing in line *k* in period *t*

$l_{v,t}$	Storage level (state of charge) for unit *v* in hour *t*
$r_{g,t}^{\mathrm{NSP}}$	Non-spinning reserve from generator *g* in hour *t*
$r_{g,t}^{\mathrm{SP}}$	Spinning reserve from generator *g* in hour *t*
$r_{v,t}$	Reserve from storage unit *v* in hour *t*
$u_{g,t}$	Commitment of generator *g* in hour *t*
$v_{g,t}$	Startup commitment of generator *g* in hour *t*
$w_{g,t}$	Shutdown commitment of generator *g* in hour *t*
$\theta_{i,t}$	Voltage angle of bus *i* in period *t*
$s_t^{\mathrm{AB}}, s_t^{\mathrm{BA}}$	Net flow on the intertie lines in one direction in period *t*
$d_{i,t}$	Demand at bus *i* for period *t*
$w_{i,t}^{\mathrm{inj}}$	Wind power injected for bus *i* in hour *t*
Π_i	Linear profit of generator *i*
Ψ_i	Net value of the demand at bus *i*
δ_i	Dual variable for fixed generator schedule, represents make whole payment for *i*
α_i	Dual variable for bounds or limits on demand *i*
$\beta_g^{\mathrm{max/min}}$	Dual variable for bounds or limits on dispatch for generator *i*
λ	Dual variable for node balance constraint (marginal price of node *i*)
λ^{DPA}	Marginal price in the new algorithm
$\lambda^{\mathrm{up/dn}}$	Conditioning increase/decrease for deviation in λ^{DPA}
u_i^{p}	Uplift payment (credit) to generation *i*
u_i^{c}	Uplift charge (debit) to generation *i*
u_i^{pd}	Uplift payment (credit) to demand *i*
u_i^{cd}	Uplift charge (debit) to demand *i*
$[\,]^*$	Optimal solution for the unit commitment problem
$[\,]^{**}$	Optimal solution for the post-unit commitment pricing problem

Parameters

a_k	Direction of line *k* relative to the network (either 1 or -1)
B_k	Line susceptance of line *k*
c_g	Linear cost of generator *g*
$c_g^{\mathrm{SU/NL}}$	Startup or no-load cost of generator *g*

c_g^{AIC}	Average incremental cost of generator *g*
c_v^{st}	OEPX cost of storage type *v* per storage cycle
$c^{\text{up/dn}}$	Cost of conditioning increase/decrease for deviation in λ^{DPA}
$C_{v,t}^{\text{max}}$	Maximum production capacity for unit *v*
$F_k^{\text{max/min}}$	Maximum or minimum line rating for line *k*
$D_{i,t}^{\text{max}}$	Demand at bus *i* for period *t*
H_t	Hurdle rate at time *t*
$I_{i,k}$	Incidence matrix for bus *i* connected to line *k*
$NTC_k^{\text{max/min}}$	Maximum or minimum net transfer capacity (line rating) for line *k*
$P_g^{\text{max/min}}$	Maximum or minimum power capacity for generator *g*
$PTDF_{i,k}$	Power transfer distribution function of node *i* and line *k*
R_g	Hourly generator ramp rate for generator *g*
R_g^{10}	Ten-minute generator ramp rate for generator *g*
$R_g^{\text{SU/SD}}$	Startup, shutdown, or five-minute ramp rates for generator *g*
$S_{v,t}^{\text{max}}$	Maximum storage capacity for unit *v*
$W_{i,t}$	Wind power forecast for bus *i* in hour *t*
$\theta^{\text{max/min}}$	Maximum or minimum angle rating
τ_g^{DT}	Minimum generator down time for generator *g*
τ_g^{UT}	Minimum generator up time for generator *g*

ACRONYMS

ac	alternating current
AIMMS	Advanced Interactive Multidimensional Modeling System
BA	balancing area
CAISO	California Independent System Operator
CCGT	combined cycle gas turbine
CH	convex hull
CHP	combined heat and power
COMPETES	Comprehensive Market Power in Electricity Transmission and Energy Simulator
DA	day-ahead
dc	direct current
DPA	dual pricing algorithm
DSM	demand-side management
ECN	Energy Research Center of the Netherlands
EIM	Energy Imbalance Market
ENTSO-E	European Network of Transmission System Operators for Electricity
EU	European Union
GAMS	General Algebraic Modeling System
GT	gas turbine
IEA	International Energy Agency
ISO	independent system operator
LMP	locational marginal price
LP	linear program
mIP	modified integer programming pricing
MIP	mixed-integer program
MISO	Midcontinent Independent System Operator
MW	Megawatt
NREL	National Renewable Energy Laboratory
NTC	net transfer capacity

OPF	optimal power flow
RT	real-time
RTO	regional transmission operator
SU, SD	startup, shutdown
UC	unit commitment
VOLL	value of lost load
WECC	Western Electricity Coordinating Council

CHAPTER 1
INTRODUCTION

Electricity is a unique and essential commodity. Unlike other commodities, electricity cannot be economically stored for long periods of time. There are no warehouses or tanks to hold it, and storage resources like batteries are not yet economically viable on a large scale. Because of this limitation, electricity demand must instantaneously be met by supply, and customer demand at a retail and wholesale level expect power delivery with high reliability. Additionally, electricity cannot be transported by any chosen path; it follows the laws of physics. These complexities pose a challenge to electric grid operators, who aim to manage the grid both reliably and at least cost. The twentieth century saw advancements in grid technology, and these advancements were honored by being called the major engineering accomplishment of that century by the National Academy of Engineering.

However, in recent decades, the electric system has not changed as drastically as other systems, such as communications. There is a common story in electricity: Alexander Graham Bell would not recognize modern telephones or the wireless network, but Thomas Edison or George Westinghouse would feel familiar with most of the components and operating procedures of modern electric systems [1]. In the last twenty years, that story is beginning to change. Wind and solar generation are fundamentally changing grid operations, advanced metering infrastructure and solid state electronics allow vastly improved state identification and control of the network, and customers are becoming more engaged in controlling their personal electric demand. Market forces are being introduced into what was formerly a vertically integrated, monopolistic industry, leading to new participants, trading arrangements, and control procedures. The combined physical and financial responsibilities of operating the electric grid are being impacted both by existing inefficiencies leftover from the grid of Edison's era and new challenges from modern technology.

This dissertation addresses some of the challenges arising from existing inefficiencies and new technologies. The chapters that follow propose methods to improve fundamental elements of power system modeling in a way that is both economically efficient and eases the integration of renewable energy or reduces emissions. The analyses are divided into four chapters, Chapters 2–5. Chapter 2 identifies an inefficient practice of allocating "lumpy" (non-convex) costs among market participants and proposes a new method to determine prices and more efficiently allocate costs. Chapter 3 suggests three improvements to reserve procurement, allocation, and activation focusing on the Dutch grid. Next, Chapter 4 examines models for integrating carbon emissions allowances into real-time operational decisions and pricing. Finally Chapter 5 compares current and proposed methods of coordinating operating decisions between regions to increase the net benefits to each. Given that costs are allocated haphazardly, present reserves are poorly quantified, carbon policies might not effectively reduce emissions, and neighboring regions fail to coordinate effectively, these chapters aim to develop methods to reduce these inefficiencies.

1.1 MOTIVATION

1.1.1 Existing Inefficiencies

Operation of electric grid has changed dramatically over the past 20 years due to restructuring of supply, transmission, and retail sales [2]. Due to the intensive capital costs of the power system, vertically integrated utilities owned and operated the generation, transmission and distribution of electricity until the 1990s [3]. Beginning in the late 1990s, the electric system in much of the U.S. and elsewhere in the world was restructured, with each segment operated by a different entity in what became known as an unbundled power system [4]. In addition, customers in some states were able to choose their own retail provider rather than using the local utility. These changes allowed for the creation of wholesale markets for electricity. Seven markets were created across the U.S., each developing under different rules over time, rules which are still undergoing modification to this day.

Pricing electricity is one issue that has been frequently reformed over time. Many markets began using a zonal model for pricing, similar to what is used in Europe today, where a large region has a single wholesale electricity price. Today, all seven organized U.S. markets use nodal pricing which distinguishes prices throughout the network given congestion, or colloquially, electric traffic. Electricity pricing has been continually reformed, with a great deal of recent attention addressing whether prices reflect the full cost of production [5]. Given the complex bid structure of power plants, academics and professionals alike are proposing new methods to price electricity. Several markets have already reformed their pricing schemes for some generators, with three others expected to propose reforms in the next year [6].

An additional contribution to inefficiencies comes from the multiple markets that have developed side-by-side in both the U.S. and Europe, because they have needed to address trade and coordination with their neighbors. Regional and national policies can make coordination difficult, since each can have different priorities and means of operation [7]. In Europe, each country operates its own real-time or balancing market, even though there are shared transmission lines between neighboring countries. In the U.S., many regions operate independently with minimal coordination with their neighbors. Operators in the Western U.S. coordinated some cross-regional trades, and in 2014, began an effort to co-optimize their real time operations under the umbrella of the Energy Imbalance Market, organized by the California

Independent System Operator [8]. These efforts to coordinate have the potential to benefit all parties if studied and enacted efficiently; however, there is no guarantee coordination in just one market will positively impact the rest. The importance and desire to coordinate has grown larger as renewable resources have entered the generation mix.

1.1.2 Inefficiencies from New Technologies

The recent growth of renewable energy has also posed many challenges for the electricity sector, which has received the attention of many policy makers, researchers, and the public [9], [10]. Power generation has been fairly predictable in system operations throughout the 20^{th} century; if a system operator requests 30 MW from a generator, that generator could produce 30 MW with high reliability. However, increased penetration into the market from renewable energy has added variability and uncertainty to the supply side. In this context, "variability" refers to volatility of the non-dispatchable net load (load minus renewable generation); "uncertainty" is defined as forecast errors or the unknown future output of net load. With uncertain forecasted generation, new market mechanisms will be necessary to ensure the electric grid is flexible enough to respond to fast changes in generation.

One method used to manage unexpected changes in generation is through operating reserves, or extra capacity held in case of unforeseen need for increased (or sometimes decreased) supply. Reserves have traditionally been held in case generation or transmission components of the electric grid go unexpectedly offline. They have not traditionally been configured to accommodate the quick changes needed when renewable energy forecasts are incorrect. This need might be met by traditional methods, but is likely to require additional means of procurement. With regional renewable integration goals, most markets are looking to ease the integration of renewable energy and curtail renewable production as infrequently as possible.

Another great challenge for market design is internalizing environmental costs, including those arising from air, water, and solid waste. A particular focus recently has been the cost of greenhouse gas emissions [2]. Many states and countries prioritize greenhouse gas emission reductions, but determining the proper method to price or penalize the emissions is not straightforward. There are many questions as to the best way to reduce total emissions and who should pay for that reduction [11]. Further complications arise as neighboring states or countries introduce different policies. The interaction of different policies may not produce expected outcomes, furthering the need for coordination and study [12].

Given the existing inefficiencies in the electric grid and the new ones brought about by renewable energy, a 1988 quote from Fred Schweppe *et al.* in *Spot Pricing of Electricity* it just as true today as it was then.

"*There is a need for fundamental changes in the ways society views electric energy.*"

([13], page xvii)

Electricity is essential for modern society; whether reading this dissertation on a screen or printed on paper, electricity was necessary for the production of the text. It is often considered an essential good, something necessary for modern life. However, most have little exposure to the production, distribution, and consumption of electricity beyond paying an electric bill. Difficulties in production and pricing are obscured from most consumers who have little notion of the existence of a wholesale market for electricity, nevertheless one operated every five minutes. In portraying a future market for electricity, Schweppe *et al.* describe a market in which both supply and demand participated. Today, supply plays an active role and demand is passive.

The quote might also suggest that power experts today should reevaluate how they view the electric grid. New technology is constantly being developed that has the potential to impact grid operations. Researchers and industry professionals alike can aid in reimagining a grid that meets the needs of all participants and ensures a sustainable future. This dissertation examines a small piece of the complex electric system and address inefficiencies therein. The main contribution of this dissertation is to suggest improvements to electricity markets in order to integrate more renewable energy into the electric system and provide the proper price signals for the market moving forward. The chapters introduce new modeling methods to make each aspect of the electric power system more efficient and provide illustrative applications. The remainder of Chapter 1 introduces the four topics in Chapters 2–5 through Sections 1.2–1.5. Section 1.2 describes the issues around wholesale electricity market pricing. Section 1.3 focuses on proposed adjustments to reserve markets. Section 1.4 probes the concerns surrounding emissions and carbon leakage in regional networks. Section 1.5 explains issues that arise from cross-regional trade and consolidation of markets in between day-ahead and real-time markets. Section 1.6 discusses the tools used within the dissertation and the scope of the remaining chapters.

1.2 ELECTRICITY MARKET PRICING

Chapter 2 focuses on electricity as a commodity, one that can be purchased and sold in a variety of contexts. Many people are most familiar with the prices they pay to a local utility on their home's monthly electric bill. These are known as retail prices for electricity, and usually regulated by a state Public Utilities Commission.[1] Although there are many options for retail rate design (e.g., time-of-use, flat rate), these prices are not determined from a competitive market auction.

Beyond retail prices, there are wholesale prices of electricity. Wholesale prices are generally not available to end-use consumers, with some exceptions for large industrial loads, such as factories. Wholesale prices can be private information, for instance, due to negotiations or bilateral contracts between a utility and a power plant, or can be publicly available. For example, there are futures prices for electricity, which result from trade on public exchanges such as the

[1] In some states there is an option for retail choice, or allowing customers to choose an alternate electricity provider. These providers are regulated, but do not need rate approval from a Public Utilities Commission.

Intercontinental Exchange (ICE). Wholesale prices are also determined through auction markets, where buyers and sellers submit bids and offers through one of the seven electricity market auctions in the U.S. These auctions are operated by nonprofit entities known as Independent System Operators (ISOs) or Regional Transmission Operators (RTOs). Six of the seven markets are regulated by the Federal Energy Regulatory Commission (FERC).[2]

By the 1936 Federal Power Act, federally-regulated markets must provide just, reasonable, and not unduly discriminatory prices for all forms of generation, including alternative energy resources [14]. Given this broad mandate, all seven markets in the U.S. have created different pricing mechanisms, although they have many similarities. Operating electricity markets and finding a price for electricity is a difficult problem both economically and mathematically. From classic Econ. 101 [15], a supply/demand graph can be drawn showing an increasing supply curve based on marginal costs and a decreasing demand curve, left graph in Figure 1-1. However, thermal generators incur both marginal and fixed operating costs in every period they operate[3]. This makes the total supply curve 'lumpy' or non-convex. With fixed costs incurred during all operational periods, it is not clear what the resulting price should be. The right graph in Figure 1-1 shows one possible visualization of a non-convex supply function, where there is a quantity under which it would be uneconomic to operate ("economic minimum operating level"), and the marginal cost of supply is a step function. This particular function happens to be quasiconvex.[4] However, not all marginal cost functions for electricity markets are quasiconvex. The convexity and quasiconvexity of the function depends on the input parameters, including whether the costs are increasing or have a quadratic term.

---

[2] The Electric Reliability Council of Texas (ERCOT) is completely contained within the state of Texas without interstate alternating current (ac) transmission lines. Because there are no interstate sales between ERCOT and another state over ac lines, it is not regulated by FERC.

[3] Fixed operating costs are not annual, long-term, or investment costs; they are costs incurred in every period that the plant operates. They are sometimes referred to as no load costs.

[4] For any two points in a quasiconvex function, all points in between will be no greater than either point; for a strict mathematical definition, see [227].

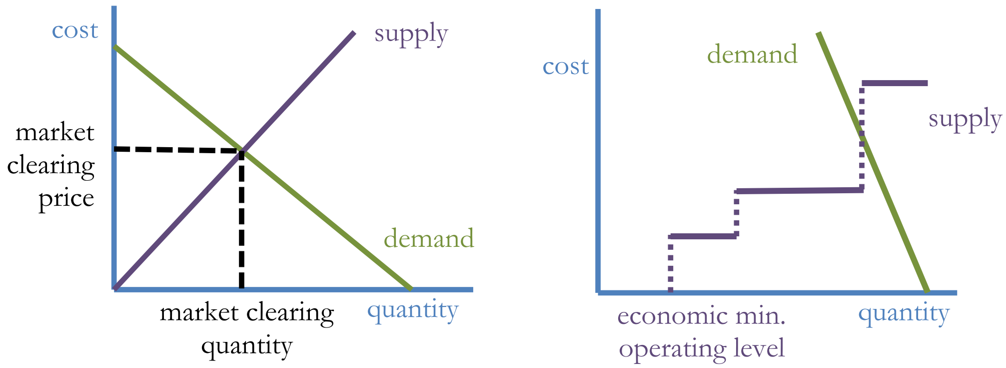

Figure 1-1 Classic supply and demand curve (left) and one visualization for a non-convex supply curve (right)

Mathematically, the optimization problem of deciding which generators to turn on and off for the following day is a difficult problem, called a mixed integer linear (or, more generally, nonlinear) problem (MILP). Unlike the classic supply demand problem, the solution to an MILP does not produce "a totally satisfactory dual," or price [16]. There are alternative possible definitions of the price that result from the dual problem and dual variables when there are binary or integer variables [17], [18]. Both economics and mathematics do not provide straightforward answers to price non-convex markets, which has created a host of proposals for electricity pricing. Due to the difficulties in non-convex markets, existing pricing proposals can result in prices that might not be non-confiscatory, revenue neutral, and incentivize efficient investments. A description of these issues, along with a new pricing mechanism is developed in Chapter 2. Examples and a discussion of desirable properties of pricing mechanisms follow the detailed mathematical description. The contribution of this chapter is a new proposal to price electricity in wholesale markets, as well as analysis of current methods and impact of market rules on pricing.

1.3 IMPROVEMENTS IN OPERATING RESERVE MARKETS

As the penetration of renewable energy grows in Europe and across the world, ease of integration is becoming a growing concern. Many European countries have set targets for renewable penetration, with an overall European goal of 20% by 2020 and 27% by 2030 [19]. As part of the European 2020 goal, The Netherlands has a binding 14% renewable energy target. Improving reserve requirements is one means to lower the cost of renewable integration and reduce wind curtailment in order to meet the 14% target.

Reserve is the additional capacity held on the system in case of errors in forecasts of load or generation, or transmission outages. Traditionally, changes to generation would entail a generator tripping offline, which could be due to a disturbance or required maintenance. Without additional backup capacity, a large loss of power might cause the system operator to 'shed load' or cut power to a portion of the system. With enough backup power online, the loss of a single generator might not cause any issues. However, renewable energy outages are different in scale and quality; for instance, they can be correlated [20]. If a large storm causes wind turbine operators across Europe to feather or lock the turbine blades (in order to ensure safety), the power lost might be higher than the loss of any one thermal generator.

There is also a more basic concern over renewable forecast uncertainty. Although forecasting has improved, German studies found that forecast error from a day-ahead forecast can be in the range of thousands of megawatts, and averages 4.5% of installed capacity [9], and can average about 20% of forecasted wind. As the penetration of renewable power increases, there will be forecast uncertainty for a large portion of generation in addition to demand. Holding additional reserves can prepare the system for such forecast errors.

Chapter 3 explores three adjustments to the current reserve procurement strategy in the Netherlands in order to reduce total operating costs and expand renewable integration. The first adjustment is changing the size of the reserve requirement to be calculated daily, as opposed to current practice where it is determined seasonally. The second is co-optimizing energy and reserve through a market, whereas current practice procures reserve through long-term contracts. Finally, the last adjustment is coordinating the allocation and activation of reserves between Northwest European countries, updating from current practice where each country acts individually. The three modifications can be made individually, in paired combinations, or all together. These combinations are simulated using a European network model with a focus on economic and renewable benefits for the Netherlands. The contribution of this chapter is the method of reserve market analysis and accompanying modeling of each type of improvement. The application also contributes an assessment of the reserve markets in the Netherlands based on economic efficiency and renewable integration.

1.4 PRICING ENVIRONMENTAL EXTERNALITIES IN REGIONS WITH ASYMMETRICAL POLICIES

Carbon emissions are a leading cause of climate change [21]. One of the major sources of carbon emissions is from power plants, primarily coal, gas, and oil fueled plants; in 2016, 1,821 million metric tons of CO_2 was emitted from the U.S. electric power sector [22]. Reducing emissions from fossil fuel power plants is a high priority for many countries and regions around the world. In the U.S., there are several states and regions that have implemented carbon reduction policies, including the Regional Greenhouse Gas Initiative in the Northeast and the cap-and-trade program in California [23]. In 2016, a landmark international agreement, the Paris Climate Agreement, aimed at reducing greenhouse gas emissions and global temperature rise was adopted by 174 countries. Although the agreement could have had an impact on national carbon reduction policies, the U.S. withdrew from the agreement in 2017 [24]. Even without the authority of the Paris Climate Agreement, existing reginal carbon policies are still in effect and will continue to influence the electric power system.

Carbon policies, such as cap-and-trade, have an impact on both planning and operational decisions of power system operators. In the case of cap-and-trade policies, power plants must hold enough allowances to cover their emissions. Those allowances can be acquired either by purchase or free allocation, depending on the exact policy demand. The allowances purchased in most existing cap-and-trade systems can impact the marginal cost offers of generators. If all states in a regional market implemented a cap-and-trade system, generators in that region could incorporate the cost of the allowances into their marginal cost bid. This would be a first-best option; the cost of carbon would directly impact dispatch and prices through generator bids into the market. However, not all states have implemented a carbon reduction mechanism. In the Eastern U.S., this has not caused many issues because of the methods states chose to implement cap-and-trade. Some are beginning to analyze options, but none have implemented market software changes to date [23], [25].

California, on the other hand, is unique because it operates a regional real-time market that extends to resources beyond California's borders. Generators in surrounding states sell power into the California market in addition to their own local utilities. California has also implemented a cap-and-trade policy through Assembly Bill 32 (AB32) [26]. Being concerned that emissions from neighboring states would increase due to capped emissions in California, the California Air Resources Board (CARB) determined that emissions from imports should be accounted for when dispatching the regional real-time market [27]. Their concerns can be categorized into carbon leakage and contract shuffling.

Carbon leakage has been defined as the increase in emissions from an unregulated region, which can be expressed as a ratio between the change in emissions in the unregulated region and the change of emissions within the regulated region [28]–[30]. This definition might not capture the accuracy of the overall reduction. For example, the cap-and-trade system might claim that total emissions inside and outside the region have been reduced by x (including an accounting for carbon associated with imports), but if total emissions in the regional electricity market are only reduced by y, and $y<x$, then $x-y$ have "leaked" out. An alternative definition is proposed in this chapter, which compares the accounted for emissions reduction within the regulated region

and from imported power subject to the regulatory system, with the actual total reduction in emissions. This alternative definition is used because some pollutant trading systems attempt to penalize estimated emissions associated with imports in an attempt to prevent leakage. In other words, leakage could be defined as just those emissions increases outside the regulated region that are unaccounted for by the regulated region. Each of the two leakage calculations provides different results and insights on the net emissions effect of alternative system dispatch models that try to limit leakage in different ways.

In California, the set of plants covered by a carbon policy is a subset of plants that participate in the real-time integrated market. Plants outside of California might be dispatched to directly or indirectly serve California load because they are less expensive, but might cause overall higher emissions. Although leakage is calculated for carbon in California within this chapter [11], emissions leakage applies broadly to other pollutants and regulation. Possible examples can include sulfur leakage in the case of the Title IV Clean Air Act SO2 trading program, where Canada or Mexico export more power to the US to replace shut-down high sulfur coal plants, and those county's sulfur emissions go up. Another example is carbon emissions in states neighboring the Northeast's Regional Greenhouse Gas Initiative (RGGI) region, where neighboring states emissions might increase due to exports while RGGI states shut down their coal-fired plants [31].

The concept of contract shuffling is a related but distinct idea that contributes to incorrect accounting of emissions associated with imports under the second definition of leakage. Contract shuffling occurs when low emitting plants outside a regulated region that would have operated even without the emissions regulations (e.g., California's AB32 cap-and-trade system) are designated as exporting to and serving the regulated region. The plants are supposedly displacing imports to the regulated region from higher emitting plants elsewhere outside the regulated region (here, California), but in fact there is no change in operations, only a change to who "provides" the imports. There is then an illusion that imports have become cleaner when it is not necessarily the case; contract shuffling is the quantification of the extent to which this occurs. While the higher emitting plants are not directly serving California customers, California's demand for imports is increasing net demand in the non-regulated region and thus contributes to the demand for the dirtier plants' output.

Both carbon leakage and contract shuffling are difficult to address in an electricity market encompassing multiple regions with different environmental rules, and these challenges are relevant to many parts of the country and world. No single answer has emerged for how to count, model, and price the externality. California has been in the lead in defining new approaches to attempt to account for leakage and contract shuffling associated with imported power, and in this Chapter, the effect of alternative approaches upon costs, emissions, leakage, profitability, and other market outcomes is simulated.

In particular, Chapter 4 analyzes five possible methods that a system operator could implement to trade emissions between regions to reduce carbon leakage and contract shuffling. The methods have trade-offs between lowering total emissions, reducing costs, and maintaining price incentives for each player. In comparing the methods, the chapter contributes to literature on pricing environmental externalities when trading partners place different value on the externality. The comparison notes the trade-offs and evaluates the strengths and weaknesses of the five methods.

1.5 TRADE BETWEEN NEIGHBORING REGIONS IN TEMPORALLY-DIFFERENTIATED MARKETS

Electricity auctions are administered in different time frames. Markets have developed in this way for many reasons. Many thermal generators have long startup times that require advanced notice. Other fossil fuel generators, especially natural gas, must secure fuel contracts the day prior to delivery. There is also uncertainty in the load forecast (customer demand), and reliability can be ensured through contingency model runs. One day prior to operation, an auction is run which determines the schedule and prices for the following twenty-fours. In some markets, one hour prior to operation another auction is run to determine advisory updates to the day-ahead schedule [32]. At least one additional auction is operated before delivery, which is fifteen or five minutes ahead of delivery to balance supply and demand given the updated forecast [4]. Often, however, modeling of electricity markets simulates a market in a single time frame rather than considering their multi-settlement nature. However, day-ahead and real-time markets can interact with each other and impact trade between regions.

Chapter 5 examines the inefficiencies with trade between neighboring regions between the day-ahead and real-time markets. The interactions between these markets might become more stressed due to renewable energy; wind and solar can be plentiful in one region and due to uncertainty with actual generation output, not be delivered in real-time. The assumptions about trade between the two markets can impact the success of renewable integration. Because of this, some markets have consolidated to simulate one large region rather than several smaller ones. In the Western U.S., California has begun an integrated real-time market (as mentioned in Section 1.4). Europe has also implemented integrated models with price coupling between regions. There are many different combinations of day-ahead and real-time integration schemes, many of which might lead to inefficiencies in the overall market even if one time frame is efficient (i.e., the day-ahead is efficient but the day-ahead and real-time together are not). This chapter simulates two regions using different trading policies in two market time frames to determine the impact of trade in the different time periods.

1.6 TOOLS AND SCOPE

Electricity markets are unique in that they combine tools from electrical engineering, operations research, and economics. Electrical engineering provides understanding of the fundamental physics of the power system, and the reliability and cost issues that might arise if operated sub-optimally. The backbone of market operations is mathematical optimization modeling from operations research. Optimization tools are used in power system planning, and day-to-day and real-time operations. While optimization models are often considered tools, algorithmic and computer science developments can become vitally important for speeding up simulation run times. Finally, economics plays a major role in both planning and operating the electric grid. The costs used in planning and bids using in operations are evaluated through financial means. Electricity markets necessarily involve participant behavior, and modeling is necessary to examine a participant's ability to exercise market power, form coalitions, and respond to incentives.

Power grid operators are also auctioneers, managing uniform price auctions every five minutes. The auctions are shaped by physical constrains but governed by economic principles. A cultural anthropologist studying electricity wrote, "…electricity alters our conventional understandings of commodities, economics, and markets" [33]. Electricity is a field that brings these three disciplines together, where engineers find themselves working in market design and economists learn the basics of power flow. The projects described in the chapters attempt to bridge the fields, bringing theory and applications from each into the different projects. The models used in each chapter are fundamentally operations research (optimization) models with economic objective functions and constraints that represent physical laws and limits and policies.

There are four projects in this dissertation, organized into Chapters 2, 3, 4, and 5. The new pricing proposal is described in Chapter 2. Pricing is introduced in Section 2.1 followed by a description of the fundamental economic principles that are the foundation of the pricing scheme in Section 2.2. Section 2.3 reviews current literature on the proposed and existing pricing schemes and evaluates them side-by-side in a table. Section 2.4 elaborates on the assumptions made, both about market rules and bidding behavior of demand. A detailed model formulation along with mathematical justification and proofs is detailed in Section 2.5. It begins with the basic unit commitment problem, its dual, and the modifications made to the dual to create the proposal, the Dual Pricing Algorithm (DPA). Formulations for other major pricing proposals are in Section 2.6. Examples using the DPA are explained in Section 2.7, along with comparison to existing pricing methods. Section 2.8 discusses some of the implications of the DPA and trends seen in the examples.

Improvements to reserve markets in the Netherlands are the focus of Chapter 3. The topic is introduced in Section 3.1, with details about how European and U.S. markets differ. Background information about European and Dutch reserve markets is found in Section 3.2, along with a review of current literature on reserves, reserve requirements, and studies on reserve coordination. The modeling framework for the reserve market is explained in Section 3.3, with details about each of the three suggested improvements. The results of the study are shown in Section 3.4, focusing on how the simulations total operating costs increase or decrease, the amount of wind curtailment, and trade between regions. The conclusions of the study are interpreted in Section 3.5, with a focus on which improvement would provide the greatest impact and the limitations of the study.

Chapter 4 evaluates the impact of greenhouse gas allowances on electricity markets. Section 4.1 introduces both topics and 4.2 provides background literature on carbon policies and details on the California cap-and-trade system. The models for greenhouse gas emissions are explained and formulated in Section 4.3. A small case study of the Western U.S. is provided in Section 4.4 with results showing the trade-offs between emissions and costs. Finally, Section 4.5 discusses the strengths and weaknesses of the five models.

The evaluation of trade and coordination between balancing areas is can be found in Chapter 5. An introduction to the issues of coordination in time between two regions can be found in Section 5.1, with a literature review in Section 5.2. Section 5.3 formulates and explains mathematical formulations for day-ahead and real-time markets. Section details the three simulations being compared for each of the day-ahead and real-time markets. The simulations results are in Section 5.6, and the conclusions about coordination and trade are assessed in Section 5.7, where total system costs are one of the main drivers of comparison.

The broad conclusions that can be drawn from the projects comprising this thesis are presented in Chapter 6. Each of the analyses suggests one or more improvements that can be made to electricity markets in the U.S. or Europe. One of the main messages of the studies is that there are significant benefits from coordination and a need for appropriate pricing. Possible future research building on these analyses is explored in in that chapter. Finally, the Appendix provides the data used for the examples in Chapter 5.

CHAPTER 2
MULTI-PERIOD DUAL PRICING ALGORITHM IN NON-CONVEX ISO MARKETS

Generation costs in electricity markets are non-convex functions of output; as a result, prices may not be monotonically non-decreasing with demand. Further, supplier revenues may not cover all variable costs. Consequently, it can be difficult to define a single price at each node that results in a balance of supply and demand. Therefore, most organized power markets in the U.S. currently pay a two-part price at each node, consisting of a public marginal price and a private make-whole payment tailored to each generator who would otherwise incur variable costs that exceed their revenue. The expense of these make-whole payments, also called uplift payments, is usually allocated evenly across all customers. This allocation method does not take into consideration who benefits from the additional costs. This paper proposes an alternative algorithm for prices in a non-convex market and a means to allocate those prices to market participants called the Dual Pricing Algorithm (DPA). Basic principles of market design are used as the foundation for the new *approach in an auction market that is revenue neutral and non-confiscatory. The general framework presents a cost allocation scheme that achieves the maximum market surplus and can be further modified to consider equity objectives defined by the system operator.*

2.1 INTRODUCTION

In many commodity markets, supply and demand curves provide a single market clearing price. In electricity markets, the non-convexities in bid and offer functions can make the traditional single market clearing price insufficient for generators to recover their variable costs (e.g., costs of start-ups) [34]. Markets in the U.S. currently provide make-whole or uplift payments for generators to ensure that they will, at a minimum, recover their operating costs.[5] Unlike the single market clearing price in markets with only marginal costs[6], supply in U.S. electricity markets can bid operational fixed costs and receive payment with a two-part price: a single price in time and space and a discriminatory private uplift payment. The system operator recovers the uplift payments in most cases by allocating it to customers, and it is often evenly distributed among consumers on a per MWh basis, even though not all consumers contribute to the need for such a payment [5]. When costs are allocated too broadly they dilute the price and location signal needed to stimulate investment in better alternatives, such as transmission infrastructure or more efficient generators. The outcome of the spot market has implications for both bilateral contracts and investment decisions. For multiple markets to be efficient, they must signal each other via public or transparent information. Meaning, prospective entrants to the market should have enough information about potential revenue sources to make an entry or investment decisions.

[5] The focus of this chapter is U.S. markets. European markets differ by region, and are generally purely financial. The market mechanism offers more complex bidding, but does not consider power flow; see [62] and [63].
[6] These are known as convex markets, defined in Section 2.2.

Day-ahead markets aim to find the surplus (consumption benefits minus costs) maximizing schedule for supply (including generator commitments) and price responsive demand. Because of non-convexities in generator costs, a mixed integer linear program called unit commitment (UC) is used.[7] In most markets, after the efficient dispatch has been determined, a pricing run determines the price of electricity at each hour and node (or bus) in the market. Pricing practice by most independent system operators (ISO) reruns the unit commitment model fixing the binary on/off decisions and relaxing the minimum operating level of the fast-start generators to zero [35].

The locational marginal prices (LMPs) result from the dual variable or shadow price of the node balance constraint. The uplift payment to a generator is determined ex-post based on its total economic loss, and varies between day-ahead and real-time (i.e., a real-time profit cannot compensate for day-ahead losses).[8] The independent system operator in the mid-Atlantic, PJM, notes that uplift average $389,000/day, but can be higher depending on conditions; during the cold weather even in January 2018, uplift costs increased to $4.3 million/day [36]. The LMPs are public and non-discriminatory, while the uplifts are discriminatory and private, lest they divulge specific generator information.[9] This public-private split means market participants and investors only know part of the information necessary to enter the market, resulting in a resulting in a weakened investment signal.

[7] Unit commitment objective functions can be difficult to solve because they are usually discontinuous. Some unit commitment formulations might result in objective functions that are quasi-convex, which are easier to solve compared to non-convex problems.

[8] As defined here, ex-post pricing means that prices are determined after the optimal unit commitment schedule has been determined; for LMP pricing, prices are based on the marginal resource's marginal cost. Prices determined before the scheduling run might not be incentive compatible for supply [230] or cause system operators to change the dispatch based on prices [231]. Ex-post prices are consistent with schedules, including any operator actions.

[9] Uplift payments are generally determined over a twenty-four hour period, and generators would be able to determine if they are receiving a payment given enough insight into their own output. In order to determine if they receive a payment, they would need to know the dispatch, bid function, and prices from the day-ahead market (including reserve market bids and quantities), and actual output and prices from the real-time market. Using this information, they would be able to calculate if they are operating at a loss; the amount of the loss is the uplift payment.

Over the last few years, many ISOs have decided to change or update aspects of their wholesale pricing mechanisms. Additionally, FERC has published a Notice of Proposed Rule Making (NOPR) on the topic of price formation. As part of the price formation process, each ISO submitted responses to a FERC Order Directing Reports requesting information on their pricing philosophy and treatment of certain variables in the unit commitment problem. Although not active at the time of submission, every ISO has determined an alternate to LMP pricing for a subset of units. These units can broadly be described as 'fast start' units, or those who can quickly startup when called upon to perform. Midcontinent ISO and ISO New England relax the binary variable in the pricing run, while PJM, New York ISO, and the CAISO[10] relax the minimum operating level and change the energy bid in the objective. These changes to pricing indicate that the LMP alone might not be able to send an incentive compatible price signal to market participants. Current implementations of fast start pricing have not necessarily changed prices significantly. MISO found that only 3% of prices changed from the baseline pricing method (LMP). This is partially due to the subset of resources that can set the price; only resources considered 'fast start' can set the price, and therefore the impact will be limited. A further discussion of price signals can be found in Section 2.2, and formulations for several alternative pricing schemes can be found in Section 2.6.

The combination of LMP and socialized uplift allocation has also caused poor investment signals. Historical examples on Cape Cod and in the upper peninsula of Michigan show that marginal pricing mechanisms can hide or misallocate funds [37], [38]. The Canal Units on Cape Cod were run daily due to their long startup times and other regional specifications. The generators only supported customers on Cape Cod, i.e., without Cape Cod demand, the generators would not be needed. However, the costs associated with operating the units were allocated to all of Lower Southeastern Massachusetts. This region as a whole did not benefit, and

[10] The CAISO mechanism is called constrained on generators (COGs), which are voluntary for generators and are not currently utilized by any generator.

it was found that costs should have been allocated primarily to Cape Cod [37]. Had the costs in Cape Cod been higher, it is possible an alternative source or upgrade could have been installed. In a similar case in the Upper Peninsula of Michigan, the Presque Isle Power Plant generated a majority of power for Michigan residents and was used for reliability in Michigan. Instead of allocating costs to Michigan, Wisconsin utilities were also charged. In the case that followed, this allocation was found unjust and unreasonable [38]. Both cases would benefit from an allocation methodology that assigns costs to responsible parties. The proposed method attempts to follow basic economic principles in order to create an efficient market.

With the underlying principle of maximizing social welfare and building on the description in [39], the proposed dual pricing algorithm outlined in this chapter aims to provide an alternate approach to efficient prices and cost allocation of make-whole payments. This paper provides a detailed explanation and justification for the single-period pricing method proposed in [39], and details the current literature on non-convex pricing. It also contributes a non-trivial extension to multiple periods and many supporting examples. The algorithm is based on the dual formulation to the post-unit commitment problem, hence it is called the Dual Pricing Algorithm. Unlike other pricing mechanisms, this algorithm allocates all costs, maintains market surplus, is non-confiscatory, and revenue neutral. These principles are examined in detail in Section 2.2. Section 2.3 discusses previous literature. Section 2.5 explains the multi-period dual pricing algorithm, and Section 2.6 describes several alternate pricing mechanisms. The results, discussion, and conclusions are in 2.7, 2.8, and 2.9.

2.2 FUNDAMENTAL ECONOMIC PRINCIPLES

2.2.1 Market Surplus and Secondary Principles

The basic principle of market design underlying the Dual Pricing Algorithm (DPA) is efficiency, as measured by the maximization of market surplus, where market surplus is the sum of consumer surplus, producer surplus (profit), and congestion surplus [40]–[42]. From this basic principle, three other guiding principles are developed below. Market scheduling software (day-ahead and real-time) attempts to determine the efficient unit commitment schedule and dispatch for resources in electric markets. Because of the non-convexities, the market clearing price is not guaranteed to cover the startup and fixed operating costs for any individual generator [43]. In order to guarantee that both generation and demand are not incentivized to leave the market (have non-negative profit and value), we include non-confiscation as the first of three secondary principles for the DPA (after the primary principle of maximizing market surplus). Non-confiscation ensures that both suppliers and demand will at least break even if they are part of the efficient dispatch; in other words, we ensure bid cost recovery. Any costs that a generator bids into the market are guaranteed to be repaid. There are exceptions to this rule in current markets, which would not be impacted by the DPA. For instance, if a participant is found to have market power or bid beyond a certain percent of a baseline, then the bid can be mitigated back to the baseline. Depending on the case, the generator might only be guaranteed recovery of the baseline bid. This principle is related to the idea of incentive compatibility and supporting prices [44]. A generator should not find it profitable to self-schedule into the market given the market prices and rules. This issue is discussed in more detail in Section 2.5.6.

The second principle is revenue neutrality, which implies revenue adequacy in the market. Specifically, we propose that the market should give out what it takes in; this applies to all energy and uplift payments from both supply and demand. The ISOs are non-profit organizations, and therefore should not plan to profit from market interactions. Any surpluses (such as congestion surpluses) are distributed either back to consumers or to holders of financial transmission rights [45], [46].

Third, the market should incentivize efficient participation and investment; this principle must hold in order to adequately build resources that will improve overall market efficiency. Efficient participation in the short term requires a pricing mechanism that supports the optimal schedule and ensures full bid cost recovery. The latter is guaranteed through the non-confiscation principle, and supporting the optimal schedule is a non-trivial assurance. In order to accomplish this goal, any mechanism that increases the price beyond marginal costs must dissuade participants from 'chasing' higher prices. Further details of the DPA's method of incentivizing participants to stay on dispatch can be found in Section 2.5.6.

Similar to incentivizing efficient dispatch, incentivizing efficient investment is a difficult task. A new entrant to the market, either generation or consumption, must consider their ability to clear the market. The uplift payment is private information; this therefore introduces an additional revenue uncertainty when a potential entrant is contemplating an investment decision. In most markets, the only public information is the price: the LMP or energy price. LMP is one valid signal for investment in a convex market, but it may be too low for a non-convex signal. A new entrant can consider whether their marginal cost bid will beat the LMP, but this information alone is not a good indication if they will be selected for the optimal dispatch. A potential plant might have a higher marginal cost, but can still clear the market with a lower fixed cost bid. Although rigorous quantification of entry and efficient investments is difficult (and outside the scope of this dissertation), it has been explored by [47]. According to [48] the investment criterion for consumers (or producers) under optimal spot pricing and convex costs is to invest in new resources if it lowers costs (or increases profits), and this occurs if and only if the investment lowers total system costs. Pricing with LMP and make-whole payments under non-convex costs could cause the investment criterion to fail for some efficient investments.

Price volatility and the presence or absence of monotonically non-decreasing price-demand relationship are two issues that arise in the non-convex pricing literature [49], [50]. Volatility often refers to rapid increases and decreases over time, which can also be the case for prices under LMP pricing. However, in this context volatility usually refers to changes in price as over increasing levels of demand. To economists, the volatility of efficient prices is of little concern. Prices should reflect the relationship between supply and demand and should include any volatility due to congestion, scarcity, or non-convexities. However, electric markets often suppress volatility in favor of 'stable' prices or fixing a scarcity problem with an out-of-market correction [51]. A pricing mechanism with a monotonically non-decreasing price-demand relationship might not reflect the true costs of the system, resulting in inefficiencies. Demand in many industries benefit from quantity discounts, or bulk purchases for a lower price. However, prices resulting from monotonically non-decreasing methods[11] can never reflect the cost savings due to higher generation production. This paper strives to create a pricing algorithm that supports market efficiency, and therefore we do not limit the method to one which will produce stable prices or prices that are monotonically non-decreasing with demand. The following two sections discuss discriminatory pricing and assumptions made in the formulation.

2.2.2 Uplift Allocation: Ramsey-Boiteux Pricing

An even distribution of uplift payments can provide misleading signals for investment. Even allocation entails splitting the needed lump payment among customers based on their total consumption (total uplift payment divided by demand in MWh). The DPA aims to allocate uplift in a discriminatory fashion among supply and demand, justifying the payments and charges with a scheme defined by Ramsey in 1927 [52]. This is often called the inverse elasticity pricing rule, or value-based pricing, because less elastic demand curves have a total higher valuation per unit, equal to the integral of the demand curve divided by quantity demanded. Its use requires

[11] Methods like convex hull pricing, described in Section 2.3.

knowledge of cost and demand functions. Ramsey's result shows that in the presence of fixed costs, the efficient result can be discriminatory pricing in proportion to demand elasticity. It can be argued that the language of the Federal Power Act would define it as not unduly discriminatory since it is efficiency enhancing.

The result was extended by Boiteux in 1956 for electricity markets, differentiating between a public and private price [53]. Ramsey-Boiteux pricing separates the single Ramsey price into a public price charged to all demand and a private discriminatory price that is different for each consumer based on that customer's elasticity; the total price paid is the sum of the two. The public price is the marginal cost (if less than average cost) and the public portion results in revenues that make up for any fixed costs not covered by the public price. Demand that is more inelastic, with a higher marginal value, will pay more of the fixed cost. Described in detail in Section 2.5, the DPA introduces uplift payments and charges for both demand and generation. In order to maintain non-confiscation for both parties, uplift is distributed according to the bids and offers placed in the market.

2.3 LITERATURE REVIEW

The literature on non-convex pricing in electricity markets can broadly be divided into proposals that advocate for a single market clearing price, and those that impose two- or multi-part pricing. U.S. markets today use multi-part pricing; a clearing price and side payments, including uplift payments. The difficulty in side-payments is determining how the market operator should recover its costs by allocating them among other market parties. Most schemes do not include specific allocation instructions, leading to inefficiencies such as the historical examples mentioned in 2.1. Alternatively, a single market clearing price is one known by all participants inside and out of the market. The price must be high enough to cover all costs of all resources in the market so as to be non-confiscatory. Given the difficulties in non-convex pricing, it is important to evaluate potential implications of new pricing methods.

O'Neill et al. provided a foundational model for two-part pricing of electricity that supports the optimal schedule [43]. The locational public price paid by all market participants at a given location is determined from the dual variable of the node balance constraint in a linear programming model of the UC problem that results from fixing the values of the binary variables at their optimal level. The second part of the price is determined from the dual variable that fixes the binary variables to their optimal schedule; a generator will only receive this payment if they suffer a loss (a negative value in the formulation in Section 2.5.1).[12] This value is the cost to cover a generator's fixed operating costs. Markets today use an approach similar to [43], with exceptions for subsets of generation, such as fast-start generators [5].

However, the O'Neill *et al.* approach suffers from having many alternative prices (degeneracy) in general. Subsequent research focused on finding prices with other desirable properties. Convex hull pricing, proposed in [54] and [55], minimizes total uplift by creating the convex hull of the total cost function so that costs are a non-decreasing convex function of load. Researchers in collaboration with the Midcontinent Independent System Operator have suggested solution techniques for the convex hull in [56], [57], and have implemented an Extended LMP (ELMP), which relaxes the binary condition on the unit commitment variable, allowing it to be between zero and one [58]. Bjørndal and Jörnsten modify the prices from [43] to create less volatile prices and uplift charges [50]. Using the same example modified in [54], they show increasing stability of average prices compared to [43].

Other models attempt to internalize uplift prices with zero-sum transfers, or payments between all participants or all generators that sum to zero. These include a "general uplift approach" using a quadratic objective in [59] and [60], and a "minimum zero-sum uplift" model that ensures that all generators break even in [49]. The method in [49] ensures that profitable generators do not to increase their profits by transferring additional payments to unprofitable

[12] A positive value from this formulation could be interpreted as a penalty assessed against a generator that is constrained off if it decides to turn on.

generators. Van Vyve proposes a non-confiscatory pricing method with separate and allocated uplift payments in [44], which results in average cost pricing if demand is inelastic. The authors in [49] also demonstrate average cost pricing for inelastic demand given two types of suppliers.

Finally, three methods attempt to create a single price that can cover both marginal and fixed costs. In [61] they use the solution technique of Lagrangian Relaxation to create a Semi-Lagrangian Relaxation, which relaxes node balance constraint and adds it to objective with a penalty price. That price is found by iterating to obtain the same objective as the original MIP and raising the clearing price to cover any fixed costs. In [62], the authors use both the primal and dual constraints to increase the clearing price to provide non-negative profits to all generators. Additional literature addressing non-convex pricing in electricity markets that does not directly suggest a new methodology can be found in [63], [64], and an additional method that solves a binary Nash game [65].

In a review article, Liberopoulos and Andrianesis analytically compare many non-convex pricing methods to determine the relative prices, payments and profits that result from each method [49]; they that find no method dominates with respect to their pricing criteria. In a similar vein to their comparison table, Table 2-1 shows a comparison of many of the methods described above. The columns show individual methodologies and the rows describe economic principles used to evaluate each method. These principles are the same as those described in Section 2.2, principles that are fundamental for the proposed pricing mechanism: maximizing market surplus, non-confiscation, revenue neutrality, and maintaining the optimal dispatch. Methods where uplift payments are determined outside of the model do not guarantee revenue adequacy (and therefore neutrality), since there might not be enough surplus from demand to pay the side-payment. All methods account for non-confiscation of supply offers; however not all explicitly account for non-confiscation of demand bids. The third row indicates whether or not the demand side was *explicitly* incorporated or if non-confiscation of demand is enforced through the pricing mechanism by itself. Many methods might be able to incorporate demand side participation in modeling but have not accounted for their participation in the publication. Transparency is designated in the fourth row. Any mechanism which includes uplift has a discriminatory and private payment; unless the payments are made public, the pricing scheme cannot be considered wholly transparent. The DPA can be adjusted or conditioned to provide either a single or two-part price, making transparency dependent on conditioning. The penultimate row describes the uplift present in the problem, whether it is allocated internally or determined after the pricing run (ex-post),[13] or zero for single part pricing. Although not documented directly in Table 2-1, all methods but the second are adjusted to incorporate fixed costs in some form into the price (to help the unit recovery costs), whereas the price resulting from [43] will be a unit's marginal cost. The last row defines a mathematical category for the pricing problem proposed, with one category defined loosely as "LP+". This category is meant to encompass math programs that can be linear, but are nontrivial to determine at each implementation. The remaining categories are linear (LP), mixed integer program (MIP), convex program (CP), and non-linear program (NLP).

Table 2-1: Comparison of Non-Convex Pricing Methods

	Two-Part Pricing							Single Price		
	Schweppe [13]	O'Neill [43]	Gribik [54], [55]	ELMP [58]	Bjørndal [50]	Galiana [59], [60]	DPA [39]	Van Vyve [44], [49]	Araoz [61]	Ruiz [62]
Maximize market surplus	Y	Y	Y	Y	Y	Y	Y	Y	Y	N
Revenue neutral	Y	N	N	N	N	Y	Y	Y	Y	Y
Includes demand side	Y	N	N	N	N	Y	Y	Y	Y	Y
Maintain optimal dispatch	Y	Y	Y	Y	Y	Y	Y	Y	Y	N
Transparency	Y	N	N	N	N	N	Y/N	Y	Y	Y
Uplifts	Ex-post	Ex-post	Ex-post	Ex-post	Ex-post	Internal	Internal	None	None	None
Pricing problem type	LP	LP	CP	LP	LP+	NLP	LP	LP	LP+	MIP*

* Combination scheduling and pricing run, linearized MINLP; all other methods are post-UC pricing runs

[13] System operators in most markets first determine the generation schedule and then determine prices through a separate 'pricing run.' Details on these two models are described in Section 2.5.1.

The table does not necessarily suggest a single dominant method, but can be used as an evaluation tool for the pricing schemes. Since no method is best in all criterion, there are trade-offs. A market designer can prioritize the criteria and select a method that best meets the needs and objectives of the market and its stakeholders.

Methods for non-convex pricing must recover any make whole payments through some allocation system. A simple approach would levy a fixed $/MWh fee to all loads. Few of the two-part pricing methods explicitly describe how uplift costs will be allocated, and there is little allocation literature focused on electricity. A general discussion of the theory and applications of cost allocation to many industries can be found in a series of essays edited by Young [66]. Electric market literature on cost allocation is mainly focused on transmission investments, such as a comparison of methods for cost allocation of transmission lines [67], [68]. As discussed in Section 2.2.2, we use Ramsey-Boiteux pricing to determine cost allocation through demand elasticity.

2.4 ASSUMPTIONS

There are several assumptions made in the model and algorithm. Each assumption is described briefly in this section, with further explanation of the impact of each in the discussion in Section 2.8.

2.4.1 Demand-Side Bidding

Although common in electricity market modeling, demand in this formulation is not infinitely valued, i.e., demand is assumed to be responsive to price. We assume that consumers bid their true value into the market; although it is possible that the bid is large (>> supply offer), it is not infinite. With completely inelastic (fixed) demand, methods like Ramsey-Boiteux cannot be used capture any flexibility for assigning payment of fixed costs, meaning there would be no advantage to discriminating among different consumers when recovering costs. Any costs that arise from the market (such as uplift) are guaranteed adequate because there is no value assigned to demand, i.e., demand is modeled as completely inelastic. In an experimental design, the authors in [69] show that demand side bidding reduces the exercise of market power and brings prices to a competitive level. Although the implementation is simple, in that we assume that demand bids are step functions, this is not a limiting assumption, as the procedure can be adapted to continuous downward sloping demand. Advances in demand side participation is not the focus of this pricing method, and well researched details on necessity, benefits, and experience can be found in Chapter 2.6 of [70].

The approach does not require an elastic demand side, but the market is more efficient when demand is elastic and demand-side participation is taken into account. Demand-side participation has become the focus of a great deal of research and interest in recent years, although the authors agree that it will take many years for the majority of demand to become elastic [71]–[73]. Mathematically, any variable the represents unserved load in market models can be considered to be a proxy for demand bids: load will not consume (be curtailed) if it reaches some value, albeit a very high value. DPA can be easily applied even if a majority of demand bids at a high value (e.g., $10,000/MWh).

Many markets today allow for price responsive demand, although participation is low. The actual percentage of load bidding into markets is difficult to estimate [74]; CAISO saw between 500-1000 MW bid in near the cap in each month in the latter half of 2016 [75]. NYISO offers economic based demand response programs, with more information available in [76].

2.4.2 Single-Node System

The DPA model is intentionally simple in order to examine the new mechanism without introducing the complications of a network. Interpreting pricing from a network model would also involve consideration of the impact of congestion. All examples in Section 2.7 are single-node or can be considered a copperplate network. The addition of an electric network and other generation constraints can be added in future research. Their addition is not necessarily straightforward, since complications arise: how should uplift be allocated across a network, can ramping constraints be properly represented, will reserve prices impact energy prices?

2.4.3 Penalties and Lost Opportunity Costs

The DPA method enforces administrative penalties rather than lost opportunity cost (LOC) payments to incentivize following the operator dispatch signal. There can be many types of lost opportunity costs in electricity markets. If generators are dispatched for reactive power support, voltage support, reserve, or other ancillary services, they are generally eligible to receive some type of lost opportunity cost payment. In these cases, the generators are asked to deviate from normal operations by the system operator and receive payments as reimbursement for gross margins they could have received if they were allowed to follow the original dispatch orders. These payments are due to operator action and, in the case of operating reserves, are automatically calculated by the market software, and are unaffected by the DPA method.

Since following the efficient dispatch along with the LMP may cause participants to forego additional profits, specific rules are required to ensure generators maintain output. As part of the DPA procedure, administrative penalties are calculated to disincentivize generators from deviating away from the dispatch signal. In addition to the penalty, any redispatch costs due to uninstructed deviations should be paid by the generator. A further discussion of penalty definitions and additional market rules is presented in Section 2.5.6.

2.5 MODEL FORMULATION AND SETTLEMENTS

In this chapter, we explain the derivation of the DPA constraints using a multi-period model, first explaining the canonical unit commitment model and dual problem. The multi-period model is an extension of the single-period model originally published in [77]. Although they are not equivalent, the following presents the derivation of the multi-period case together with some observations on the single-period model. Section 2.5.1 explains a basic unit commitment model and its dual problem. The settlement that is obtained from the basic model based on [43] is explained in Section 2.5.2. Modifications made to the dual problem to formulate the DPA are described in Section 2.5.3, and additional constraints to manage non-unique prices are in Section 2.5.4. The full DPA formulation is in Section 2.5.5, with additional rules explained in 2.5.6.

2.5.1 Unit Commitment Model and Dual

The formulation in (2-1)-(2-7) is the canonical unit commitment problem from [43]. Without transmission, reserve, and other generator characteristics, the unit commitment problem becomes a simple model that, as the name suggests, commits generators for the following day and used in real-time commitment for short start units. The objective in (2-1) is to maximize social welfare or market surplus. Both demand bids and generator offers are included, where generators can bid startup and fixed operating costs. Constraint (2-2) matches supply and demand in each time period. The generator minimum and maximum operating limits are in (2-3), and (2-4) defines the logic for the startup variable. Constraint (2-5) places bounds on the maximum and minimum dispatch for demand. Finally, (2-6) and (2-7) define the variables u_{it} and z_{it} to be binary, making this model a mixed-integer linear program. Chapters 3 and 4 will explain a more complex unit commitment model, which also includes generator ramping, minimum up and down times, transmission, and reserve constraints.

$$\max \sum_{t\in T}\left(\sum_{i\in D} b_{it}d_{it} - \sum_{i\in G}\left(c_{it}p_{it} + c_{it}^{OC}u_{it} + c_i^{SU}z_{it}\right)\right) \tag{2-1}$$

$$\sum_{i\in D} d_{it} - \sum_{i\in G} p_{it} = 0 \qquad \forall t \in T \tag{2-2}$$

$$p_i^{min}u_{it} \le p_{it} \le p_i^{max}u_{it} \qquad \forall i \in G, t \in T \tag{2-3}$$

$$u_{it} - u_{i,t-1} \le z_{it} \qquad \forall i \in G, t \in \{2..T\} \tag{2-4}$$

$$0 \le d_{it} \le d_i^{max} \qquad \forall i \in D, t \in T \tag{2-5}$$

$$u_{it} \in \{0,1\} \qquad \forall i \in G, t \in T \tag{2-6}$$

$$z_{it} \in \{0,1\} \qquad \forall i \in G, t \in T \tag{2-7}$$

The formulation in (2-8)-(2-14) is the post-unit commitment pricing run. The post-unit commitment model is the second model the market software runs, where the scheduling run in (2-1)-(2-7) is the first. There are many other runs that can occur, including models for market power mitigation, where supplier offers can be modified if they are suspected of attempting to exercise market power, and runs to ensure reliability. The literature of mixed-integer programming has proposed several alternative definitions of dual problems for MILPs, so a determination must be made about how to define prices. In all U.S. markets for slow generation, as in [43], marginal prices can be calculated from the LP that arises when the binary variables are fixed to their optimal solution from the scheduling run. The major differences in this model compared to the scheduling run are constraints (2-13)-(2-14), which fix the binary variables (u_{it} and z_{it}) to the optimal value from the unit commitment problem. The variables in the right column are the dual variables for each constraint. Variations on this pricing method used in actual markets are described in Section 2.6.

$$\max \textstyle\sum_{t\in T}(\sum_{i\in D} b_{it}d_{it} - \sum_{i\in G}(c_{it}p_{it} + c_{it}^{OC}u_{it} + c_i^{SU}z_{it})) \tag{2-8}$$

$$\sum_{i\in D} d_{it} - \sum_{i\in G} p_{it} = 0 \qquad \forall t \in T \qquad \lambda_t \tag{2-9}$$

$$p_i^{min}u_{it} \le p_{it} \le p_i^{max}u_{it} \qquad \forall i \in G, t \in T \qquad \beta_{it}^{max}, \beta_{it}^{min} \tag{2-10}$$

$$u_{it} - u_{i,t-1} \le z_{it} \qquad \forall i \in G, t \in \{2..T\} \qquad \delta_{it}^{SU} \tag{2-11}$$

$$0 \le d_{it} \le d_i^{max} \qquad \forall i \in D, t \in T \qquad \alpha_{it}^{max} \tag{2-12}$$

$$u_{it} = u_{it}^{*} \qquad \forall i \in G, t \in T \qquad \delta_{it}^{u} \tag{2-13}$$

$$z_{it} = z_{it}^{*} \qquad \forall i \in G, t \in T \qquad \delta_{it}^{z} \tag{2-14}$$

The dual formulation of the pricing run can be found in constraints (2-15)-(2-20). Parallel to the above model, the right column shows the primal variables associated with each constraint in the dual problem.

$$\min \sum_{t\in T}(\sum_{i\in D} d_i^{max}\alpha_{it}^{max} - \sum_{i\in G}(u_{it}^*\delta_{it}^u + z_{it}^*\delta_{it}^z)) \tag{2-15}$$

$$\lambda_t + \alpha_{it}^{max} \geq b_{it} \qquad \forall i \in D, t \in T \qquad d_{it} \tag{2-16}$$

$$-\lambda_t + \beta_{it}^{max} - \beta_{it}^{min} \geq -c_{it} \qquad \forall i \in G, t \in T \qquad p_{it} \tag{2-17}$$

$$\delta_{it}^{SU} - \delta_{i,t+1}^{SU} + \delta_{it}^u - p_i^{max}\beta_{it}^{max} + p_i^{min}\beta_{it}^{min} = -c_{it}^{OC} \quad \forall i \in G, t \in \{1..T-1\} \qquad u_{it} \tag{2-18}$$

$$\delta_{it}^z - \delta_{it}^{SU} = -c_{it}^{SU} \qquad \forall i \in G, t \in T \qquad z_{it} \tag{2-19}$$

$$\alpha_{it}^{max}, \beta_{it}^{max}, \beta_{it}^{min} \geq 0 \qquad \forall i \in G \cup D, t \in T \tag{2-20}$$

2.5.2 Settlements Derived from the Dual Problem

Using the dual formulation, we can formulate the dual pricing algorithm using the economic principles discussed in Section 2.2. From strong duality[14] of the primal and dual post-unit commitment linear programs, the optimal primal and dual solutions (if both are feasible) must satisfy:

$$\sum_{t\in T}(\sum_{i\in D} b_{it}d_{it} - \sum_{i\in G}(c_{it}p_{it} + c_{it}^{OC}u_{it} + c_i^{SU}z_{it})) \tag{2-21}$$
$$= \sum_{t\in T}(\sum_{i\in D} d_i^{max}\alpha_{it}^{max} - \sum_{i\in G}(u_{it}^*\delta_{it}^u + z_{it}^*\delta_{it}^z)).$$

For generators, the dual constraints for the dispatch and commitment variables in (2-17) and (2-18) must be modified to reflect the startup decision in order to enforce non-confiscation in the pricing method. In current markets, the price paid by the ISO is λ_t^{**}, or the dual variable of the node balance constraint. For $i \in G$, from (2-17) and complementary slackness,

$$\left(-\lambda_t^{**} + \beta_{it}^{max**} - \beta_{it}^{min**} + c_{it}\right)p_{it}^* = 0 \qquad \forall i \in G, t \in T \tag{2-22}$$

If $u_{it}^* = 1$, then from (2-10) and complementary slackness,

$$(p_{it}^* - p_i^{max})\beta_{it}^{max**} = 0 \qquad \forall i \in G^+, t \in T \tag{2-23}$$

$$\left(p_{it}^* - p_i^{min}\right)\beta_{it}^{min**} = 0 \qquad \forall i \in G^+, t \in T \tag{2-24}$$

Using (2-22), (2-23), and (2-24) in (2-18), a one period model (excluding δ_{it}^{SU}) produces the classic economic result that profits are revenue less costs, or

---

[14] Strong duality occurs when both the primal and the dual problems have optimal solutions, then their objective functions will be equal [17].

$$\delta_{it}^{u**} = p_{it}^{*}(\lambda_t^{**} - c_{it}) - c_{it}^{OC} \qquad \forall i \in G. \qquad (2\text{-}25)$$

That is, δ_{it}^{u**} is the LMP payment less the marginal and fixed costs incurred. Unfortunately, there is no guarantee that δ_{it}^{u*} is non-negative, that is, non-confiscatory. In the multi-period case, we sum together the δ_{it}^{u**} constraints in (2-18) for all periods to create (2-26) and sum the startup periods of (2-19) to obtain constraint (2-27). The sum of (2-26) and (2-27) then define a total linear profit function for the operating periods in (2-28). Both (2-23) and (2-24) are again substituted into (2-18), producing a total profit function for a multi-period model in (2-28), where τ_i is the number of time periods in which the generator starts up ($t \in T | z_{it} = 1$), T^r is a dynamic set that refers to the total run periods (all startups to shutdowns, $T^r = \{t \in T | u_{it} = 1\}$).

$$\Delta_i^u = \sum_{t \in T^r} \delta_{it}^{u**} = \sum_{t \in T^r} \left(p_{it}^{*}(\lambda_t^{**} - c_{it}) - c_{it}^{OC}\right) \qquad \forall i \in G^+ \qquad (2\text{-}26)$$

$$\Delta_i^z = \sum_{t \in T} \delta_{it}^{z**} = -\tau_i c_i^{SU} \qquad \forall i \in G^+ \qquad (2\text{-}27)$$

$$\begin{aligned} \Pi_i &= \Delta_i^u + \Delta_i^z \\ &= \sum_{t \in T^r} (p_{it}^{*}(\lambda_t^{**} - c_{it}) - c_{it}^{OC}) - \tau_i c_i^{SU} \end{aligned} \qquad \forall i \in G^+ \qquad (2\text{-}28)$$

Similar to the single period result, (2-28) provides the classic economic result that the total profit under LMP pricing for generation is the payment received less the variable and fixed costs.

In the cases when the generator is turned off, or $u_{it}^{*} = 0$, we can write a separate profit condition; rearranging (2-17), we have

$$\beta_{it}^{max**} \geq \lambda_t^{**} - c_{it} \qquad \forall i \in G^0, t \in T. \qquad (2\text{-}29)$$

Substituting (2-29) into the startup condition of (2-18), and summing over time as described in (2-28), we obtain

$$\Pi_i = \sum_{t \in T^r} (p_{it}^{max}(\lambda_t^{**} - c_{it}) - c_{it}^{OC}) - \tau_i c_i^{SU} \qquad \forall i \in G^0 \qquad (2\text{-}30)$$

We have the following four potential outcomes for u_{it}^{*} that demonstrate the need for make-whole payment and penalties.

- When $u_{it}^{*} = 1$ in one or more periods and $\Delta_i^u + \Delta_i^z < 0$, then a make-whole payment, $-(\Delta_i^u + \Delta_i^z)$, in addition to the LMP payment, which does not cover the offered cost, is needed to avoid confiscation.
- For periods when $u_{it}^{*} = 1$ or when $u_{it}^{*} = 0$ in one or more periods and $\Delta_i^u + \Delta_i^z > 0$, then the LMP and a penalty or an LOC payment is needed to incentivize generators to stay on dispatch. The LMP with penalty provides enough disincentive to price-chasing and self-scheduling behavior.
- When $u_{it}^{*} = 0, \forall t$ (the generator is not dispatch in any period) and $\Delta_i^u + \Delta_i^z \leq 0$, then the LMP sends the correct price signal. The generator would not profit from an LMP payment.

As described in [43], there are two basic pieces of the settlement: the price (λ_t^{**}) and the make-whole payment ($\Delta_i^u + \Delta_i^z$). Without make-whole payments, the results from the post-unit commitment problem can be confiscatory. Due to the non-convexities in the market, there is also a third part of the settlement, which can be imposed as a penalty or a payment. If $\lambda_t^{**} > c_{it}$ and $p_{it}^{*} < p_i^{\max}$ or $\lambda_t^{**} > c_{it}$[15] and $u_{it}^{*} = 0$ then the generator faces a lost opportunity cost; if it had been online or been dispatched to its full capacity, it could have made a profit. In order to discourage a generator from changing its output or committing themselves in the next period or market (expecting the price to stay the same), a payment or penalty can be imposed that is the cost of the lost opportunity. We will assume here that the penalty for self-scheduling is high enough to prevent inefficient dispatch and the dispatch signal is a quantity signal. The penalty would be at minimum the value of the lost opportunity cost.

This subsection describes the basic pricing scheme from [43], with an extension to multiple periods. The result is non-confiscatory for supply but does not specify a make-whole payment

[15] A full valuation would consider the average incremental cost of operating, not necessarily the marginal cost alone, since fixed operating costs would not be recovered if the generator chooses to self-dispatch or commit themselves into the market. However, this evaluation is dependent on the generator and their behavior and expectations.

allocation method that is non-confiscatory for demand. The next subsection modifies into the primal and dual constraints to uphold the economic principles from Section 2.2 (including non-confiscation) to construct the Dual Pricing Algorithm. Note that there are no side payments for demand, because there are no binary variables for demand bids; however, if future demand bids consist of binary variables, the formulation can be adapted to include any fixed operating costs incurred.

2.5.3 Transforming the Dual Problem into the Dual Pricing Algorithm

In this subsection, the Dual Pricing Algorithm is developed from the primal and dual constraints for the post-unit commitment problem. We now add constraints to the equilibrium conditions on the market-clearing quantity to reallocate the market surplus and ensure non-confiscation of supply and demand. Consequently, the new price, λ_t^{DPA}, is no longer necessarily the LMP. In order to ensure non-confiscation, new uplift payments and charges are introduced in the algorithm. Both supply and demand can be charged uplift, and similarly can be paid uplift. The market operator announces the price and settles the uplift payments and charges with each market participant, both generators and consumers.

The DPA scheme guarantees non-confiscation of generator supply offers. We demonstrate above that the profit as defined in (2-25) and (2-28) can be negative. To ensure non-confiscation in the DPA, we introduce an uplift payment, μ_{it}^{p}, and uplift charge, μ_{it}^{c}, that can be impose for each market participant. We can redefine the profit condition in and (2-28) with non-confiscation of the profits Π_i as

$$\Pi_i = \sum\nolimits_{t \in T_r} \left(p_{it}^{*}\left(\lambda_t^{DPA} - c_{it} + \mu_{it}^{p} - \mu_{it}^{c}\right) - c_i^{OC} \right) - \tau_i c_i^{SU} \qquad \forall i \in G \qquad \text{(2-31)}$$

where we substitute λ_t^{DPA} for λ_t^{**} and introduce the uplift quantity $p_{it}^{*}\left(\mu_{it}^{p} - \mu_{it}^{c}\right)$ in each period. For supply, the profit or non-confiscation condition now is

$$\Pi_i \geq 0 \qquad \forall i \in G^{+} \qquad \text{(2-32)}$$

In order for the DPA to guarantee non-confiscation of demand bids, we substitute $(\mu_{it}^{pd} - \mu_{it}^{cd})$ for α_{it}^{max**} and λ_t^{DPA} for λ_t^{**} in the complementary slackness condition from (2-16), i.e.,

$$d_{it}^{*}\left(b_{it} - \lambda_t^{DPA} + \mu_{it}^{pd} - \mu_{it}^{cd}\right) = 0 \qquad \forall i \in D, t \in T \qquad (2\text{-}33)$$

This relationship is then summed over the commitment period to account for all payments and charges. The net value, $d_{it}^{*}b_{it} - d_{it}^{*}\lambda_t^{**}$, for $d_{it}^{*} > 0$ can be defined as

$$\Psi_i = \sum\nolimits_{t \in T_r} d_{it}^{*}\left(b_{it} - \lambda_t^{DPA} + \mu_{it}^{pd} - \mu_{it}^{cd}\right) \qquad \forall i \in D \qquad (2\text{-}34)$$

The net value must be nonnegative to ensure non-confiscation, enforced by

$$\Psi_i \geq 0 \qquad \forall i \in D^{+} \qquad (2\text{-}35)$$

Since the market surplus is positive, the uplift payments and charges simply reallocate market surplus. Uplift payments and charges are participant specific, avoiding confiscation of any one participant. Without discriminatory uplift pricing, make-whole payments recovered uniformly across demand could result in confiscation.

For demand bids $i \in D^0$ not selected (i.e., $d_{it}^{*} = 0$), the net profit is zero, $\Psi_i = 0$. This is true for any feasible solution to the non-convex post-unit commitment market model. Substituting λ_t^{DPA} for λ_t^{**} in (2-16) and setting $\alpha_{it}^{max} = 0$ because $d_{it}^{*} = 0$, we obtain the following lower bound on the DPA energy price.

$$\lambda_t^{DPA} \geq b_{it} \qquad \forall i \in D^0, t \in T \qquad (2\text{-}36)$$

This implies the new λ_t^{DPA} will be high enough to ensure an out-of-market bid will not consume. In other words, unserved load will prefer not to self-dispatch or take recourse actions that will lower the market surplus.

Since the DPA scheme is run post unit commitment, the criteria to maintain market surplus is already satisfied for the optimal solution to $(p_{it}^{*}, d_{it}^{*}, u_{it}^{*}, z_{it}^{*})$; therefore constraint (2-37) is enforced but redundant in the formulation.

$$\sum\nolimits_{i \in G} \Pi_i + \sum\nolimits_{i \in D} \Psi_i = MS^{*} \qquad (2\text{-}37)$$

To ensure revenue adequacy, we balance the uplift payments and uplift charges through the following revenue neutrality condition. As noted for other pricing methods in [49], this type of constraint is called a "zero-sum uplift" constraint.

$$\sum_{t \in T} \left[\sum_{i \in D^+} d_i^* \left(\mu_{it}^{pd} - \mu_{it}^{cd} \right) + \sum_{i \in G^+} p_i^* \left(\mu_{it}^{p} - \mu_{it}^{c} \right) \right] = 0 \tag{2-38}$$

2.5.4 Conditioning

The prices that result from the DPA are in general non-unique. Therefore, a choice can be made from the alternative prices based on the preference of the operator and market participants. The price has no coefficient in the objective function, allowing it to vary through a range of prices (shown in the example in 2.7.1.b). Specific allocation criteria can be embedded in the model and produce conditioning such as perceived equity or increased transparency. Conditioning or tuning are broad terms that refer to the additional constraints added to the DPA formulation to create unique prices or modify the price to resemble operator preference. If the region wishes to keep prices close to the dispatch LMP, the DPA can maintain close prices and additionally allocate the uplift. If the market operators prefer a single market clearing price with no uplift payments, the algorithm can be modified to determine a single price. By providing tuning capabilities, we acknowledge that factors outside of the mathematical formulation or economic theory can drive decision making.

There are several types of constraints that can be employed to condition the price. The ones proposed here use surplus and slack variables to push the DPA price close to the LMP and produce a unique price; however, a different base price (other than the LMP) can be used if preferred by the system operator. The constraints can either assign the surplus and slack variables to individual periods or use a single set of variables for all periods. The impact of these constraints can be seen in examples in Section 2.7.

We offer two options to 'tune' the price. Both condition the LMP by keeping the new price, λ_t^{DPA}, close to the dispatch LMP, λ_t^*, with penalties for deviations. To minimize the relative deviation, we construct (2-40), and (2-45) in the next section. In (2-40) there is a single surplus and slack variable for the entire run time, while in (2-45) there are surplus and slack variables for each period. Using these constraints can either concentrate high prices in one period or distribute them across all periods. The second option in (2-39) minimizes the absolute deviation across time.

$$\lambda_t^{\mathrm{DPA}} - \lambda_t^{**} - \lambda_t^{up} + \lambda_t^{dn} = 0 \qquad \forall\, t \in T \tag{2-39}$$

$$(\lambda_t^{DPA} - \lambda_t^{**})/\lambda_t^{**} - \lambda^{up} + \lambda^{dn} = 0 \qquad \forall\, t \in T \tag{2-40}$$

If an operator was concerned about price spikes, we also may want to condition the uplift payments. Many possibilities can be considered. One possibility is to limit the maximum allowable payment and charge by the constraints listed in (2-41) and (2-42). However, this may result in insufficient cost allocation (an infeasible solution).

$$\mu_{it}^{p} \le \mu_i^{p\mathrm{max}}, \mu_{it}^{c} \le \mu_i^{c\mathrm{max}} \qquad \forall i \in G^+, t \in T \tag{2-41}$$

$$\mu_{it}^{pd} \le \mu_i^{pd\mathrm{max}}, \mu_{it}^{cd} \le \mu_i^{cd\mathrm{max}} \qquad \forall i \in D^+, t \in T \tag{2-42}$$

2.5.5 Dual Pricing Algorithm (DPA) Formulation

We now formulate the DPA model using the modifications of the dual problem described in Section 2.5.1, with one conditioning constraint to make the prices unique. The objective in (2-43) minimizes the uplift payments made by both generation and demand. Constraint (2-44) is the uplift revenue neutrality constraint and (2-45) is one option for price conditioning. Constraints (2-46) and (2-47) are the generation profit and demand value respectively. A demand bid lower bound on pricing is in (2-48), and (2-49) and (2-50) ensure non-confiscation. Finally, (2-51) contains the nonnegative variables.

$$\min \sum_{t \in T} \left[\sum_{i \in D^+} d_{it}^* \mu_{it}^{pd} + \sum_{i \in G^+} p_{it}^* \mu_{it}^{p} + c^{up} \lambda_t^{up} + c^{dn} \lambda_t^{dn} \right] \tag{2-43}$$

$$\sum_{t \in T}\left[\sum_{i \in D^+} d_{it}^*\left(\mu_{it}^{pd} - \mu_{it}^{cd}\right) + \sum_{i \in G^+} p_{it}^*\left(\mu_{it}^{p} - \mu_{it}^{c}\right)\right] = 0 \qquad (2\text{-}44)$$

$$(\lambda_t^{DPA} - \lambda_t^{**})/\lambda_t^{**} - \lambda_t^{up} + \lambda_t^{dn} = 0 \qquad \forall t \in T \qquad (2\text{-}45)$$

$$\Psi_i = \sum\nolimits_{t \in T_r} d_{it}^*\left(b_{it} - \lambda_t^{DPA} + \mu_{it}^{pd} - \mu_{it}^{cd}\right) \qquad \forall i \in D^+ \qquad (2\text{-}46)$$

$$\Pi_i = \sum\nolimits_{t \in T_r}\left(p_{it}^*\left(\lambda_t^{DPA} - c_{it} + \mu_{it}^{p} - \mu_{it}^{c}\right) - u_{it}^* c_{it}^{OC}\right) - \tau c_i^{SU} \qquad \forall i \in G^+ \qquad (2\text{-}47)$$

$$\lambda_t^{DPA} \geq b_{it} \qquad \forall i \in D^0, t \in T \qquad (2\text{-}48)$$

$$\Psi_i \geq 0 \qquad \forall i \in D^+ \qquad (2\text{-}49)$$

$$\Pi_i \geq 0 \qquad \forall i \in G^+ \qquad (2\text{-}50)$$

$$\mu_{it}^{p}, \mu_{it}^{c}, \mu_{it}^{pd}, \mu_{it}^{cd}, \lambda_t^{up}, \lambda_t^{dn} \geq 0 \qquad \forall i \in D \cup G, t \in T \qquad (2\text{-}51)$$

Theorem 1. If there exists an optimal solution to the primal unit commitment problem, that is, the maximize market surplus problem, then there is a feasible solution to DPA.

Proof. A feasible solution to (1) is obtained with $d_{it} = p_{it} = z_{it} = 0$ and $MS = 0$. From the post-unit commitment problem and summing together (2-44), (2-46), and (2-47), we have $MS^* = \sum_{i \in G} \Pi_i + \sum_{i \in D} \Psi_i \geq 0$.

From complementary slackness of (2-3) with the binary fixed at its optimal value, there are three possible cases for the values of $p_{it}^*, \beta_{it}^{max**}$, and β_{it}^{min**}:

(a) if $p_{it}^* = p_i^{max}$, then $\beta_{it}^{max**} > 0$ and $\beta_{it}^{min**} = 0$;

(b) if $p_{it}^* = p_i^{min}$, then $\beta_{it}^{max**} = 0$ and $\beta_{it}^{min**} > 0$;

(c) if $p_{it}^* \in \left(p_i^{min}, p_i^{max}\right)$, then $\beta_{it}^{max**} = \beta_{it}^{min**} = 0$.

Therefore $p_{it}^*\left(\beta_{it}^{max**} - \beta_{it}^{min**}\right) = p_{it}^{max}\beta_{it}^{max**} - p_{it}^{min}\beta_{it}^{min**}$. From complementary slackness of (2c), $p_{it}^*(\lambda_t^{**} - c_{it}) = p_{it}^*\left(\beta_{it}^{max**} - \beta_{it}^{min**}\right)$. As shown in (2-28), $\Delta_i^u + \Delta_i^z$ is the linear surplus of generator $i \in G$. From complementary slackness of (2-16), $d_{it}^*(b_{it} - \lambda_t^{**}) = d_{it}^*\alpha_{it}^{max**}$, and $d_{it}^*\alpha_{it}^{max**} \geq 0$ since d_{it}^* and α_{it}^{max**} are both nonnegative.

We partition $i \in G$ into three sets $G' = \{i \in G: \Delta_i^u + \Delta_i^z \geq 0 \text{ and } u_{iT_r}^* = 1\}$, $G'' = \{i \in G: \Delta_i^u + \Delta_i^z < 0 \text{ and } u_{iT_r}^* = 1\}$, and $G''' = \{i \in G: u_{it}^* = 0\}$. $\Pi_i = 0$ for all $i \in G'''$.

$MS^* = \sum_{i\in G} \Pi_i + \sum_{i \in D} \Psi_i =$

$\sum_{i\in G'} \Pi_i + \sum_{i\in G''} \Pi_i + \sum_{i\in D} \Psi_i \geq 0$

$\sum_{i\in G'} \Pi_i + \sum_{i\in D} \Psi_i \geq -\sum_{i\in G''} \Pi_i$

Let $\lambda_t^{\mathrm{DPA}} = \lambda_t^{**}$ and use the previously mentioned complementary slackness conditions to see that

$\Pi_i = \Delta_i^u + \Delta_i^z + \sum_{t\in T}\left(p_{it}^* \mu_{it}^p - p_{it}^* \mu_{it}^c\right)$ and

$\Psi_i = \sum_{t\in T}\ d_{it}^* \alpha_{it}^{max**} + d_{it}^* \mu_{it}^{pd} - d_{it}^* \mu_{it}^{cd}$.

Let $\mu_{it}^p = \mu_{it}^{pd} = 0$ on the LHS, $\mu_{it}^c = 0$ on the RHS, and substituting for Π_i and Ψ_i:

$\sum_{i\in G'}(\Delta_i^u + \Delta_i^z - \sum_{t\in T} p_{it}^* \mu_{it}^c) + \sum_{i\in D,T} d_{it}^* \alpha_{it}^{max**} - d_{it}^* \mu_{it}^{cd}$

$\geq -\sum_{i\in G''} \Delta_i^u + \Delta_i^z + \sum_{t\in T} p_{it}^* \mu_{it}^p$

Then we can select payments and charges:

For all $i \in G''$, let $\sum_{t\in T} p_{it}^* \mu_{it}^p = -(\Delta_i^u + \Delta_i^z) > 0$. This satisfies (2-48) for $i \in G''$.

For all $i \in G'$ select μ_{it}^c and μ_{it}^{cd} such that,

$\sum_{t\in T} \sum_{i\in G'} p_i^* \mu_i^c + \sum_{i\in D} d_i^* \mu_i^{cd} = \sum_{t\in T} \sum_{i\in G''} p_{it}^* \mu_{it}^p$,

$\sum_{t\in T} p_{it}^* \mu_{it}^c \leq \Delta_i^u + \Delta_i^z$ for $i \in G'$, and

$\sum_{t\in T} d_{it}^* \mu_{it}^{cd} \leq \sum_{t\in T} d_{it}^* \alpha_{it}^{max**}$ for $i \in D$.

The selection criteria are equivalent to (2-44), (2-49), and (2-50).

From strong duality of the post-unit commitment problem and its dual problem, we know that

$MS^* = \sum_{i\in G} \Delta_i^u + \Delta_i^z + \sum_{i\in D, t\in T} d_{it}^* \alpha_{it}^{max**} \geq 0$, and

$\sum_{i\in G'} \Delta_i^u + \Delta_i^z + \sum_{i\in D, t\in T} d_{it}^* \alpha_{it}^{max**} \geq -\sum_{i\in G''} \delta_{iT_r}^{u**} + \delta_{iT_1}^{z**}$.

Therefore, the generators $i \in G'$ and demands $i \in D$ have enough linear surplus to satisfy (2-49) and (2-50).

Constraints (2-44), (2-46), (2-47), and (2-51) are satisfied by the construction.

For $d_{it}^{*} = 0$, complementary slackness requires $\lambda_t^{**} \geq b_{it}$, so (2-48) is satisfied and the DPA has a feasible solution. ■

2.5.6 Rules for Uninstructed Deviations

A new pricing rule cannot be implemented in isolation. Consideration of generator and demand behavior must be explicitly incorporated into the price or developed alongside market rules. Given the chance, all market participants will act in their own best interest and not that of the social optimum. When a generator sees a price above its marginal costs, it will be inclined to produce as much as possible to increase profits. Similarly, rational customers will want to consume more given lower prices. The latter issue is addressed in constraint (2-48). The former leads to a discussion of uninstructed deviations, or colloquially, price chasing, and methods needed to maintain the efficient dispatch.

There are two main methods utilized to ensure market participants stay on dispatch: lost opportunity costs and penalties defined by market rules. The use of either should make a participant indifferent between deviating from the dispatch signal and maintaining the optimal dispatch. For ease of understanding, the following simple example will illustrate the issue. Two generators and two loads have the characteristics in Table 2-2.

.

2.5.6.a Lost Opportunity Costs: Demonstrative Example

Unit commitment would dispatch Generator A to 30 MW and Generator B to 70 MW. The LMP would be \$40/MWh, which is the cost of Generator A. Both Generators A and B would receive uplift payments, \$500 and \$1900 respectively. To examine the impact of opportunity costs and penalties, we first compare the outcome using an existing pricing method, namely, Midcontinent ISO's ELMP. This method relaxes the binary variables in (2-6) and (2-7) rather than fixing them to the optimal level as was done in (2-13) and (2-14). In the problem above we can relax the binary constraints and determine the single period ELMP, \$65/MWh.

Table 2-2: Characteristics of example market participants

Participant	Marginal cost or value ($/MWh)	Startup cost ($)	Min Capacity (MW)	Max Capacity (MW)
Generator A	40	500	0	98
Generator B	60	500	70	100
Demand 1	1000	-	0	80
Demand 2	67	-	0	20

Under ELMP, Generator A would be paid a price of $65/MWh, much higher than its marginal costs, and yet it is not dispatched to its maximum. The generator would understandably be incentivized to produce at its maximum capacity to capture further profits. Any pricing method that raises the price above the traditional LMP will need to designate a method to disincentivize this behavior. If Generator A deviated from the dispatch signal, it would make an additional $1575. The operator can either pay Generator A $1575 or penalize Generator A $1575 if it deviates. Both actions should have the same impact, either should incentivize Generator A to maintain the 30 MW dispatch signal.

Assuming an even distribution of the opportunity cost payment, demand would be required to pay an additional $15.75/MWh. Due to the high marginal bid of Demand 1, it will see a small decrease in total value. Demand 2 will be asked to pay more than its marginal bid (payment of $80.75 > $67). Although Demand 1 would still maintain positive surplus, Demand 2 suffers a loss. While there is enough consumer surplus to support the prices, Demand 2 does not have enough value to pay the new price. The method of paying a lost opportunity cost is confiscatory if uplift is socialized.

If the operator instead chose to impose a penalty with a value of $1575, the Generator would not see benefit in increasing output and the market would remain revenue adequate. Penalties also do not impact confiscation of demand. This is a stylized example, and changes to the bids and offers would necessarily change the outcome. This simple example explains part of the motivation for using penalties; there are cases in which paying lost opportunity costs are not revenue adequate and result in confiscation. New proposals for pricing must consider all impacts to the market; both penalties and lost opportunity costs can result in the same incentive for generators, but the latter does not ensure bid cost recovery for all participants.

One of the reasons the DPA uses penalties over lost opportunity costs is to ensure non-confiscation and revenue adequacy. The example above showed confiscation, while the example in Table 2-3 shows revenue inadequacy.

Table 2-3 Characteristics of example market participants

Buyer	Value ($/MWh)	Max demand (MW)	Gen	Marginal cost ($/MWh)	Startup cost ($)	Min Capacity (MW)	Max Capacity (MW)
1	45	60	A	30	900	0	200
			B	40	100	10	200

With the parameters given in Table 2-3, the basic unit commitment market would clear with Generator B supplying all demand for Buyer 1, and Generator A remaining off. This occurs because of the high startup costs for Generator A. The market clearing price is $40/MWh (due to Generator B), and the market surplus is $200. Because the market clearing price is above the marginal cost of Generator A, it will be incentivized to startup unless a penalty is imposed or a lost opportunity cost is paid. The penalty or cost should be greater than $1100 (= 200 MWh * ($40/MWh - $30/MWh) - $900); however, there is not enough market surplus to pay Generator A. Therefore, the market is revenue inadequate if a lost opportunity cost is used. Again, this is a simple example, but shows the undesirable characteristics of lost opportunity costs.

In the case of completely inelastic demand, the market will always be revenue adequate because all costs can be levied on demand. Any uplift or lost opportunity cost will be paid through the 'unlimited' demand side value. This proposal models a two-sided market to show that this is not the case if the market is elastic, even if it is only partially elastic. While this is also true for uplift payments, the following example shows how lost opportunity costs are not guaranteed to be revenue adequate. There is not always enough surplus in the market to cover an additional lost opportunity cost payment.

2.5.6.b Lost Opportunity Costs: Current Practice

There are many types of lost opportunity costs currently used in electricity markets. If generators are dispatched for reactive power support, voltage support, reserve, or other ancillary services, they are generally eligible to receive lost opportunity cost payments. In these cases, the generators are asked to deviate from normal operations by the system operator, and receive payments as reimbursement for money they could have received if they were allowed to follow the original dispatch orders. These payments are due to operator action, and are different than those discussed in Section 2.5.6.a.

We are specifically discussing (active) generator deviations from the optimal dispatch signal sent by the operator; meaning it is an intentional deviation often called "uninstructed deviations." Most markets discourage this practice through penalties. At publication, only ISO-NE pays lost opportunity costs. All other ISOs choose to penalize generator deviations beyond a dead-band [35], [78]–[82]. NYISO confiscates the energy, meaning they do not pay for the additional energy produced. MISO already has thresholds for 'uninstructed deviations' and is considering updating the rules and penalties based on suggestions from their market monitor. In the 2016 State of the Market Report, the market monitor suggests that improving these penalties "will improve suppliers' incentives to follow MISO's dispatch signals and will, in turn, improve reliability and lower overall system costs" [82]. In addition to confirming the argument presented in 2.5.6.a, the references above show that the use of penalties in the DPA is a continuation of most ISOs' current practice and not a significant change to current electricity markets.

2.5.6.c Lost Opportunity Costs: Clarification of Methodology and Liquidated Damages

Any auction market follows a set of rules. If a participant chooses to enter the market, they must agree to follow and conform to the rules. Each electricity market operator establishes rules in their tariff. In order to bid and be cleared in the market, these rules must be followed. If a rule states that a participant must follow the dispatch signal for the periods in which it actively bids into and is cleared in the market, then it should not need additional incentives to follow the rule. Rather, there should be consequences for failing to follow the policies. The consequences we propose consist of a penalty and payment for liquidated damages.

The example above demonstrated one example penalty, and this section will further define the quantitative method proposed to impose penalties. The basic procedure uses the value of the lost opportunity cost payment as a penalty. Any additional profit that a generator would make by deviating away from the optimal dispatch should be the total value of the penalty. A straightforward calculation is

$$\sum_{t\in T} \lambda_t^{DPA}(P_i^{\max} - p_{it}^*). \tag{2-52}$$

For the quantity of energy generated above dispatch signal, the generator would receive the LMP rather than the DPA. In the case of the example above, if Generator A deviated and produced an additional 10 MW, the 10 MW would be compensated at $40/MWh. In addition to the penalty and the lower price, we contend that any generator that deviates also pay the redispatch costs, or liquidated damages.

Generally, contracts that involve the future exchange of money or the promise of performance have a liquidated damages stipulation. A liquidated damages payment is monetary compensation for a loss from a breach of contract. This stipulation establishes what must be paid if a party fails to perform as promised. In contract law, liquidated damages are considered reasonable in light of the anticipated or actual harm caused by the nonperformance (breach of contract). In the case of electricity markets, the liquidated damages payment would be the rebalance costs due to a generator's deviation from their optimal schedule. The deviation would be a breach of contract, and the harm to the system is the rebalance costs.

In total, the DPA will be composed of several elements and accompanying rules. The energy price and uplift payment or charge is the primary result of the DPA formulation. To discourage uninstructed generator deviations, the excess power generated will only receive the LMP as an energy price. In addition, there will be a liquidated damages stipulation requiring the generator pay any rebalance costs. The combination of prices and market rules should incentivize a core economic principle from Section 2.2: efficient participation and investment.

2.6 ALTERNATIVE PRICING METHODS

2.6.1 Extended LMP

As discussed in Section 2.3, there are many alternative proposals for pricing methodology. While most ISOs have implemented some form of alternative pricing for fast-start units, the most prominent is called Extended LMP, implemented by Midcontinent ISO with similar methods used by New England ISO [35], [58], [83]. Due to the difficulties implementing convex hull pricing, the convex hull is approximated by relaxing the binary commitment variables. The method is under its second update (ELMP 2.0), and therefore the formulation shown below is a single period approximation, not an exact replicate of current practice. The examples in Section 2.7 compare DPA prices to ELMP, therefore we offer a brief explanation of the method. Constraints (2-53)–(2-57) are the same as the unit commitment and post-unit commitment problem. The difference lies in the relaxation of the binary variables in (2-58)–(2-59). By relaxing the commitment, the minimum capacity of the generator can dip below p_i^{min}, and fixed costs can be incorporated into the price. The new price will minimize uplift and lost opportunity cost payments, but not eliminate uplift or ensure non-confiscation [84].

$$\max \sum_{t\in T}\left(\sum_{i\in D} b_{it}d_{it} - \sum_{i\in G}(c_{it}p_{it} + c_{it}^{OC}u_{it} + c_i^{SU}z_{it})\right) \tag{2-53}$$

$$\sum_{i\in D} d_{it} - \sum_{i\in G} p_{it} = 0 \qquad \forall t \in T \qquad \lambda_t \tag{2-54}$$

$$p_i^{min}u_{it} \le p_{it} \le p_i^{max}u_{it} \qquad \forall i \in G, t \in T \qquad \beta_{it}^{max}, \beta_{it}^{min} \tag{2-55}$$

$$u_{it} - u_{i,t-1} \le z_{it} \qquad \forall i \in G, t \in \{2..T\} \qquad \delta_{it}^{SU} \tag{2-56}$$

$$0 \le d_{it} \le d_i^{max} \qquad \forall i \in D, t \in T \qquad \alpha_{it}^{max} \tag{2-57}$$

$$0 \le u_{it} \le 1 \qquad \forall i \in G, t \in T \qquad \delta_{it}^{u} \tag{2-58}$$

$$0 \le z_{it} \le 1 \qquad \forall i \in G, t \in T \qquad \delta_{it}^{z} \qquad (2\text{-}59)$$

2.6.2 Average Incremental Cost Pricing

The minimum single price to recuperate all fixed operating costs is the average incremental cost of the marginal unit. Any price below the marginal unit's average incremental cost will result in an uplift payment. Although average incremental cost pricing is not implemented in any market, it is a useful method for comparison because it results in no uplift. It is also helpful to compare the prices from average incremental cost pricing because the DPA often produces equivalent prices. There is no rigorous proof for this statement, primarily because it is an observation from working examples and not necessarily always equivalent. Unlike other pricing methods which will necessarily produce certain prices (see Table 2-4), the prices resulting from the DPA uphold the economic principles explained in Section 2.2 without a constant direct correlation to the cost function of the marginal generator. In a simple model with a single period, single node, and no other generation characteristics (i.e., ramping, reserve, etc.), the LMP will be the marginal cost of the marginal generator.

Table 2-4 Characteristics from Other Pricing Methods

	LMP	**ELMP**	**AIC**
Single-period	c_{it}	$c_{it} + \left(\frac{c_{it}^{\mathrm{OC}}}{P_i^{\max}}\right)$	$c_{it} + \left(\frac{c_{it}^{\mathrm{OC}}}{p_{it}^{*}}\right)$
Multi-period	c_{it}	$c_{it} + \left(\frac{c_{it}^{\mathrm{OC}} + c_i^{\mathrm{SU}}/MRT}{P_i^{\max}}\right)$	$c_{it} + \left(\frac{c_{it}^{\mathrm{OC}} + c_i^{\mathrm{SU}} / \sum_{\forall t} p_{it}^{*}}{p_{it}^{*}}\right)$

Note: The allocation of startup costs for AIC pricing depends on the implementation of multi-period rules

2.7 EXAMPLES

We explore several examples to illustrate the capability and flexibility of the DPA. The first section shows examples for a single-period model, demonstrating some fundamental concepts for DPA pricing and comparing with popular examples in the literature. The next section shows the prices that can come from multi-period models. These examples show how pricing in a single period context does not always easily extend to multiple periods, meaning multi-period pricing can be more complex and is not single periods added together.

2.7.1 Single Bus, Single Time Period Examples

2.7.1.a Two-Generator Example

Considering a one period example with the data provided in Table 2-5, the optimal dispatch is Generator A = 40 MW and Generator B = 90 MW with an LMP of \$60/MWh. Generator A will profit while Generator B will operate at a \$500 loss, shown in Table 2-7. The DPA determines the modified LMP to be λ^{DPA} = \$65.56/MWh, which recovers Generator B's startup costs by charging an additional \$5.56/MWh to both buyers. Since the new LMP is above the bid of Buyer 2, Buyer 1 pays an additional \$1.37/MWh (μ_1^{pd}) making Buyer 2 break even by receiving a payment of \$4.56/MWh (μ_2^{cd}). The λ^{DPA} reflects the incremental cost of serving load and the resulting settlement leaves Generator B and Buyer 2 at a break-even point, Generator A with increased profits, and Buyer 1 with decreased additional value.

These prices and profits can be compared with the ELMP model. As shown in Table 2-7, ELMP produces a price of \$62.50, which is the marginal cost plus the startup cost amortized over the total capacity. Generator A receives more profit than under LMP pricing, but less compared to DPA pricing. Generator B requires a make-whole payment, although it is smaller than the payment under LMP.

Table 2-5 Generator Costs

Gen	Marginal cost ($/MWh)	Startup cost ($)	P^{min} (MW)	P^{max} (MW)
A	40	500	0	40
B	60	500	10	200

Table 2-6 Demand Function

Buyer	Value ($)	Max demand (MW)
1	100	100
2	61	30

Table 2-7 Total Profit and Value

	Traditional LMP	DPA	ELMP
Price	$60/MWh	$65.56/MWh	$62.50/MWh
Generator A Profit	+$300	+$522.22	+$400
Generator B Profit	**–$500**	**$0**	**–$275**
Buyer 1 Value	+$4,000	+$3367	+$3750
Buyer 2 Value	+$30	$0	+$45
Uplift	$500	$76.67	$275

2.7.1.b MISO Four Generator Example

Using data from a MISO sample problem in [83], we can look at a range of demand levels for a single period problem. The generator costs are in Table 2-8, and the single demand has a value of $100/MWh. Figure 2-1 shows the clearing price for three different pricing methods for demand from 0 MW – 350 MW. The dispatch LMP (λ^*) spikes when moving from the cheaper generators to the more expensive (A to C), since it must turn on the generator with the lowest minimum and highest marginal cost (D) to match demand. The ELMP is monotonically non-decreasing with demand, forming steps when the expensive generator is needed. Similar to the dispatch LMP, the DPA price also spikes when the expensive generator is dispatched. The prices then decrease, showing quantity discounts as the generator reaches it maximum. There are no uplift payments needed with the DPA, while there are payments required from the LMP and ELMP. The prices without condition would range from the lower-bound in Figure 2-1 to $100/MWh, the demand offer. Since λ^{DPA} has no coefficient in the objective, the prices output can vary within that range in every simulation; the GAMS solver returned a value of $100/MWh without conditioning, although other solvers might return the lower-bound value.

Table 2-8 Generator Costs

Gen	Marginal cost ($/MWh)	Startup cost ($)	P^{min} (MW)	P^{max} (MW)
A	50	500	20	100
B	52	500	20	100
C	55	500	20	100
D	65	40	5	50

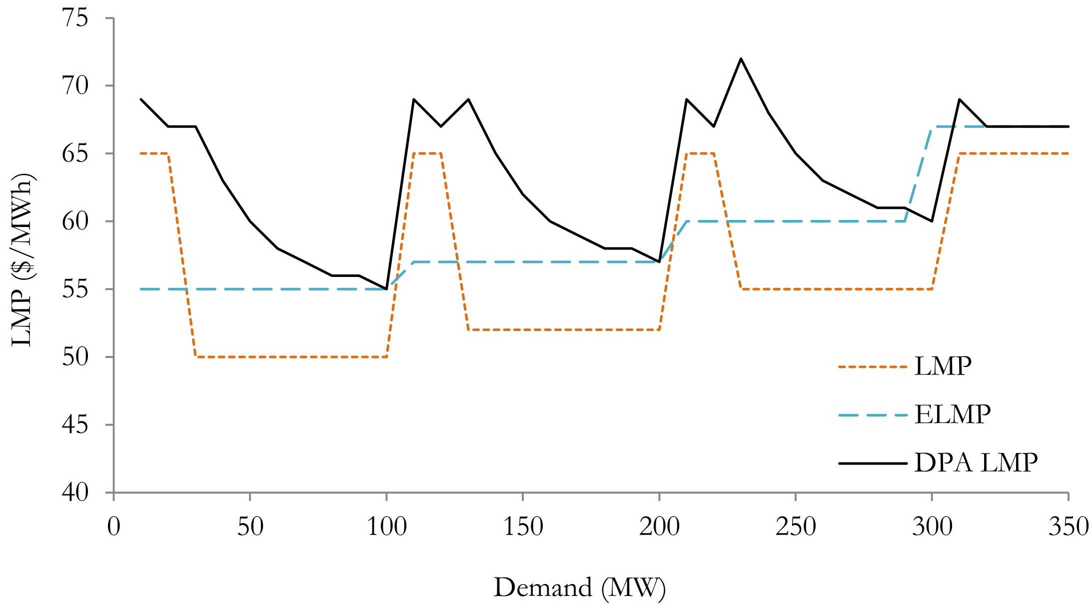

Figure 2-1 Snapshots of different demand levels with three different pricing methods

2.7.1.c Scarf Example

We simulated other small test examples with similar results. A benchmark example created by Scarf in [85] has been used to demonstrate the versatility of pricing methods. The DPA is compared with a traditional LMP and uplift in Figure 2-2. The figure shows the changes in price as demand quantity increases. The prices and resulting uplift payments are shown with blue solid and dashed lines, while the DPA prices and uplift are shown in black and orange. There are no uplift payments made at any demand level, and prices oscillate between $6/MWh and $7/MWh under DPA pricing, the latter being the price of the generator with a high marginal cost and no startup cost. Under LMP pricing, uplift is required and prices oscillate between $3/MWh and $7/MWh.

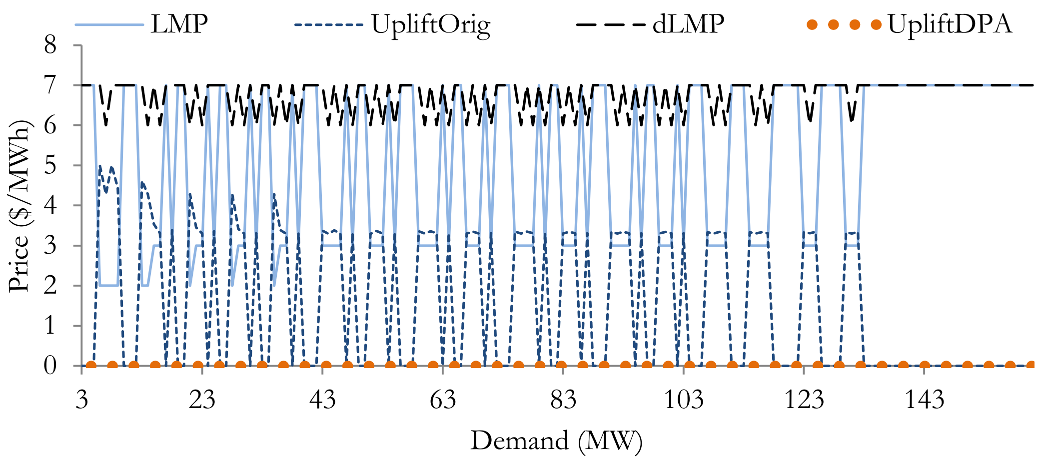

Figure 2-2 Prices resulting from the modified Scarf example compared with the traditional method of determining prices and uplift payments

2.7.2 Multi-Period Examples

The single period examples demonstrate the prices across varying demand levels, while the following multi-period examples focus on the prices and payments across time.

2.7.2.a Multi-Period, Multi-Generator Example with Conditioning

The generator characteristics for this set of examples are found in Table 2-9 and the demand value and quantity, and reserve data are found in Table 2-10.

Table 2-9 Generator Data

Gen	Marginal cost ($/MWh)	Startup cost ($)	No Load cost ($/h)	P^{min} (MW)	P^{max} (MW)
A	30	900	100	200	1200
B	50	600	100	50	80
C	60	360	100	25	50

Table 2-10 Hourly Data

	1	2	3	4	5	6	7	8
Demand 1, Value $200/MWh	510	528	546	573	582	588	594	564
Demand 2, Value $80/MWh	340	352	364	382	388	392	396	376
Reserve (MW)	85	20	20	20	20	20	20	20

The resulting prices are in Table 2-11, showing the difference between the two types of DPA conditioning and the traditional LMP, λ_t^*. In the first conditioned DPA price, λ_t^{DPA}, the penalties are imposed in every period (constraint (2-45)), while $\lambda_t^{\text{DPA}\prime}$ imposes a single penalty on all periods (constraint (2-40)). The impact is uniform prices for all periods compared to a higher price in a single period. Another notable impact is on uplift payments. The dispatch LMP imposes a $1700 uplift payment on the system, while both conditioned DPA prices incur no uplift. This is an example where the DPA produces a single market clearing price, and there is no need to follow uplift allocation guidelines.

Table 2-11 Prices & Payments

	1	2	3	4	5	6	7	8	Uplift ($)
λ_t^*	30	30	30	30	30	30	30	30	1700
λ_t^{DPA}	30.23	30.23	30.23	30.23	30.23	30.23	30.23	30.23	0
$\lambda_t^{\text{DPA}\prime}$	30	30	30	30	30	30	31.72	30	0

2.7.2.b RTS Test Case

We examine the generation from a modified single zone RTS96 test case [86]. The generator characteristics and load data are found in Table 2-12 and Figure 2-3. All generators were located at a single node with 24 hourly simulations. The resulting generator profits and demand value are in Table 2-13, and prices in Figure 2-4. As expected, the total social welfare remains the same between the two simulations. In order to reduce uplift and provide proper incentives for investment, there is a transfer of surplus between consumers and producers. With zero uplift, the price provides a transparent indicator for investment; it allows investors to evaluate if their unit could enter the dispatch. While the only guaranteed method of analysis for market entry would involve rerunning the dispatch with the potential unit, the transparency of the DPA price sends more information to potential entrants than a marginal pricing method alone.

Table 2-12 Generator Data

Gen	Quantity	P^{max} (MW)	P^{min} (MW)	$c^{startup}$ ($)	$c^{marginal}$ ($/MWh)	c^{noload} ($/h)
Oil/CT	4	20	15.8	76	163	1139
Coal/Steam	4	76	15.2	1061	19.64	131
Oil/Steam	3	100	25	4754	75.64	840
Oil/Steam	3	197	68.95	6510	74.75	1160
Oil/Steam	5	12	2.4	571	94.74	73
Coal/Steam	2	155	54.25	1696	15.46	253
Nuclear	2	400	100	2400	5.46	215

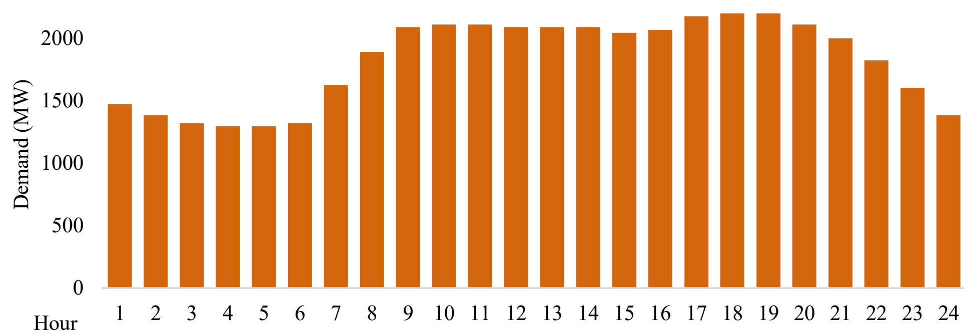

Figure 2-3 Hourly demand for the modified RTS example

Table 2-13 Surplus & Payments

	Traditional LMP	DPA
Generator Profits	$2,244,014	$4,765,784
Consumer Value	$5,233,475	$2,711,704
Uplift	$10,768	$0

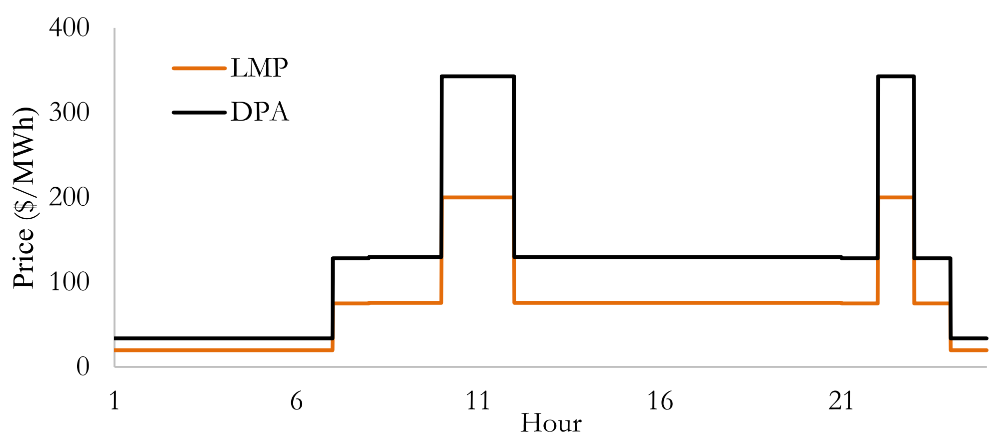

Figure 2-4 Hourly price comparison

2.7.2.c Multi-Period Comparison with Uplift

This 8-period example compares prices that result from four different pricing models: LMP, ELMP, DPA, and average incremental cost or locational incremental price (LIP). The generator characteristics can be found in Table 2-8 and demand data in Table 2-14. The dispatch can be found in Table 2-15, where the underlined values represent generators dispatched to their minimum operating levels. The resulting energy prices in Figure 2-5 show several trends. Both LMP and DPA remain the same for the first four periods, while ELMP and LIP increase for periods 3 and 4. In these periods, Gen B is at its minimum capacity, making Gen A the marginal unit. The same occurs when Gen C is at its minimum in period 6; both the ELMP and LIP increase, while the LMP and DPA remain the same. The DPA behaves this way due to its tuning to the LMP. In this case, the conditioning constraint in (2-40) is modified so that the surplus and slack variables are the same in every period, i.e., $(\lambda_t^{DPA} - \lambda_t^{**})/\lambda_t^{**} - \lambda^{up} + \lambda^{dn} = 0$. The DPA is slightly higher than the LMP in each period, with a small allocated uplift. The DPA allocates the uplift payment of \$8.89/MWh to Gen C and charges Demand 2 \$0.49/MWh. Demand 2 has a higher value bid compared to Demand 1, and receives the full uplift charge. The uplift for the other methods can be calculated as a single lump-sum charge to demand; the LMP uplift payment is \$3110, ELMP payment is \$197.20, and LIP payment is \$0.

Table 2-14 Hourly Data

	1	2	3	4	5	6	7	8
Demand 1, Value \$80/MWh	510	550	620	597	600	636	640	520
Demand 2, Value \$200/MWh	600	608	626	653	662	668	674	644

Table 2-15 Hourly Generation in MWh where Underlined Values are at Minimum Capacity

	1	2	3	4	5	6	7	8
Gen A	1110	1136	1196	1200	1200	1200	1200	1164
Gen B	0	0	<u>50</u>	<u>50</u>	62	79	80	0
Gen C	0	0	0	0	0	<u>25</u>	34	0

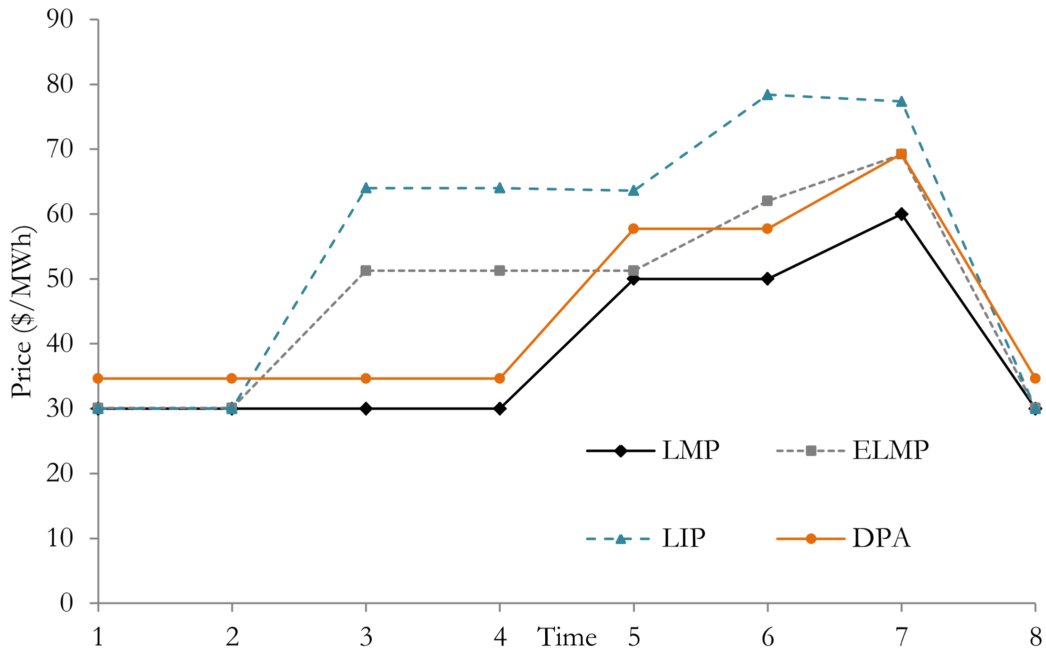

Figure 2-5 Prices from 8-period comparison example

2.8 DISCUSSION

The examples in the previous section show the prices and costs that can result from the DPA. While not guaranteed to always occur, there are several common trends in the examples. DPA prices tend to be higher than the traditional LMP and the ELMP. This is not surprising due to the incorporation of fixed costs, and low or no uplift payments. While ELMP minimizes uplift payments, they are often nonzero. With zero uplift, the additional cost is incorporated into the prices, causing them to be generally higher than LMP prices. Higher prices should not be perceived as positive or negative; however, when there are no private side payments, there is increased transparency in the market. Additionally, compared to pricing mechanisms that are non-decreasing (like convex hull), prices are more volatile. Due to fixed costs, DPA prices are closer to the average incremental cost of delivering power, which is a decreasing function with respect to demand for each generator. As discussed in Section 2.2, volatility should not be considered an objectionable trait, rather one that can reveal the true value of producing power.

Pricing mechanisms should produce efficient prices, ones that support the optimal schedule. In combination with deviation penalties, the DPA prices support the optimal schedule and recover all parts of both generation bids and demand offers. The prices also signal points of entry into the market. With low or no uplift payments, new entrants can better evaluate if their incremental costs are below the clearing price. While this is not a guaranteed point of entry, it provides more market information than the traditional LMP. When side payments are needed, the algorithm allocates them to both supply and demand in particular periods. The endogenous allocation ensures that demand does not pay more than its offer and supply is made whole. In a one-sided market (where demand in inelastic), demand will pay any price and revenue adequacy is guaranteed. Even in markets today there are elastic bids, a trend that is likely to increase as markets change in response to the shifting resource mix.

The DPA method and resulting prices share similarities with existing methods discussed in Section 2.3. Similar to all methods but [62], the DPA maintains the optimal market surplus. It fixes the optimal schedule, like the formulation in [43] and also fixes the dispatch level. The revenue neutrality characteristic is shared by the single price methods in [44], [61], [62], and the additional uplift revenue neutrality condition is shared by [59], [60]. Uplift calculations are done endogenously in the DPA and [59], [60], while other methods either have zero uplift or calculate the payment ex-post. Both [44] and the DPA may result in prices that are the marginal generator's average incremental cost. The DPA method strives to support basic economic market principles and formulating a pricing method around these principles will cause overlap with existing methods. However, the DPA upholds all principles simultaneously while incorporating demand side participation and multiple periods. The extension to multiple periods helps allocate costs to the periods that cause them. Some methods might have direct extensions to multiple periods, but most cases are not well-defined.

Unlike other pricing methods which produce prices with a consistent relationship to the cost function of the marginal cost generator, the DPA prices uphold the economic principles from

2.2 without a strict relationship to the marginal generator's cost function. The prices often result in the average incremental cost. This price has the benefit of eliminating uplift payments, therefore ensuring bid cost recovery. As mentioned in 2.5.6, the high prices that can occur must be supported by strong market rules to dissuade price chasing and support the efficient dispatch. While the prices will be higher than the LMP, the average incremental cost is an incremental price signal for non-convex entry, whereas the LMP is a weak signal.

The DPA introduces both supply-side and demand-side payments in order to recover costs due to prices that are above or below a resource's costs. At present, the demand-side does not pay uplift based on elasticity, and introducing such payments might adjust demand behavior. In theory, it might act like a pay-as-bid scheme: demand with low elasticity pays demand with high elasticity, both recovering only their bid-in value. A pay-as-bid scheme for generation can increases prices, causing generation to bid strategically at the estimated price rather than their true marginal cost [87]. Demand might act in the same way, meaning, they would lower their bid to pay less for electricity. Since there is very little demand-side bidding in markets today, it is difficult to evaluate this theory on actual customer behavior. In order to actively bid at low values, the demand must be willing to stop consumption if not selected in the optimal dispatch (due to low bids). This is a risky endeavor, and would depend on the type of consumer. It is unlikely that all consumers will be able to stop consuming; however, it also places a burden on inelastic demand to pay a greater amount of uplift. If the penetration of price responsive demand increases, further analysis must be done to assess their ability to strategically bid. In this case, additional rules or regulations might be created.

2.9 CONCLUSION

Spot prices should provide proper incentives for both operations and investment. Electricity is unlike other commodities due to the fixed costs necessarily incurred during operation and the need to physically balance supply and demand. Due to these non-convexities, it is difficult to determine the 'right' price for electricity. Methods suggested in the literature often consider only one aspect of pricing or are contingent on inelastic demand. This project proposes the Dual Pricing Algorithm, which brings together many principles surrounding pricing mechanisms: maximizing market surplus, revenue neutrality, non-confiscation, transparency, signals for market entry, and side payment allocation. The DPA is an ex-post pricing scheme that upholds these principles and can be adapted to particular system operator needs. It is a linear program, making it computationally efficient, and can be incorporated into current ISO software. The approach is applied to multiple time horizons and can easily include additional operational constraints, e.g., reserve requirements. Further work can be done to incorporate these constraints and evaluate the algorithm on a large network model.

CHAPTER 3
EFFICIENT ACQUISITION OF GENERATION RESERVES TO BACK-UP WIND IN DUTCH ELECTRICITY MARKETS

Reserve capacity is needed in an electric system in case of a contingency. With the increased penetration of renewable energy in the Netherlands, its variable output can require additional reserve on the system. The Netherlands currently procures reserve months in advance through long time contracts with little coordination with its neighbors. If additional reserve is needed, it should be procured at least cost to the system. This project suggests improvements that can be made to the process of procuring, allocating, and activating reserve. The simulations use a multiple time scales and modeling across many regions. A model of Europe called COMPETES is used to analyze the improvements, employing a future scenario with a high penetration of wind across countries. Comparison of different improvements will elucidate which have the most benefits to the system as a whole, and the methodology can be applied in many regions worldwide.

3.1 INTRODUCTION

As discussed in the introduction to this dissertation, electricity is a unique commodity. Supply must meet demand instantaneously, with no economic opportunities for large-scale storage in most parts of the world. Due to this instantaneous need, grid operators have traditionally held backup power capacity at the ready in case of emergencies or contingencies. Theses can range from load and renewable energy forecast errors to outages of large generators or transmission lines. If there is more than expected load on the grid, power plants have automatic detection that increases or decreases output depending on the frequency or regulation signal. A large change to forecasted load or the loss of a generator would require more power than an automatic response can produce. In that case, operators call online generators to increase output. The additional capacity that is held at the ready in case of a contingency or forecast error is broadly called operating reserves [88], the technical aspects of which will be discussed in the following section. Operating reserves are primarily procured from generators online or those that can come online quickly (fast start generators). Procurement and use between the U.S. and Europe varies greatly, where U.S. uses a market to procure reserve and Europe uses long term contracts [89].

As variable resources have come online, the need for additional capacity has expanded beyond generator outages and load forecast error. Renewable forecast error is present in both the day-ahead and balancing (real-time) markets, forcing operators and policymakers to consider how to handle the additional variability and uncertainty. Reserve requirements can be inflated to reflect the added difficulty of renewable forecast error, but this change alone is likely not enough to manage the quality and quantity of renewable power coming online [88]. The magnitude of power can be large in some areas, causing strain on other plants to ramp up to support large renewable generation swings. In other regions, the reserve needed might necessitate a different type of resource, such as a change in reactive power output. While there are many hypothetical examples, the issue remains that operators must evaluate how to change reserve products in light of increased renewable energy penetration. This chapter evaluates alternative improvements that could be made to reserve procurement in order to increase efficiency in the network. The improvements look holistically at reserve, from sizing to allocation to activation, to determine what changes will decrease costs and impact renewable curtailment.[16] The Netherlands case study will demonstrate which improvements make the largest impact, enabling policy makers to focus on the most cost-effective strategies. The analysis is holistic, combining multiple aspects of reserve through different markets (day-ahead and balancing) and across regions.

Reserves can be used in many stages of power system operation. The loss of a generator during operations is a short-term problem. The need for reserve also arises in long-term planning; if an operator builds a future system to meet demand exactly, there might be periods of shortfall. Instead, planning reserve determines the amount of extra capacity needed for investment decisions, which can impact the reliability of the grid [90]. Although it can have an impact on future reliability [91], the focus of this chapter is on operational reserve for short-term operations.

Obtaining operational reserve can generally be divided into three steps: sizing, allocation, and activation [92]. The sizing of reserve determines how much is needed to be procured, called a reserve requirement; for instance, three percent of load forecast is needed in every hour or a fixed amount of MW, as seen in Figure 3-1. The required reserve is then allocated to particular generators either through long term contracts or through a mechanism in an energy market. In the real-time or balancing time-frame, the reserves are activated or consumed as needed. This activation might be the result of operator action, or through market mechanisms (for instance, in the way the flexible ramping is deployed through normal energy dispatch in MISO and CAISO). This phase assesses whether there were enough reserves procured and if they were deliverable

[16] An alternative approach would be use of stochastic programming, which determines reserves endogenously rather than ex ante. Although it has not been implemented in an actual market setting, research is ongoing, with seminal papers [232]–[234]. With increased uncertainty from renewable generation, stochastic modeling can become increasingly important.

when needed. While there is great focus on the first stage, sizing, the next two are equally important for the reliability of the system. An abundance of reserve procured in a congested area that cannot deliver during a contingency will cause an outage, just as under procurement would.[17]

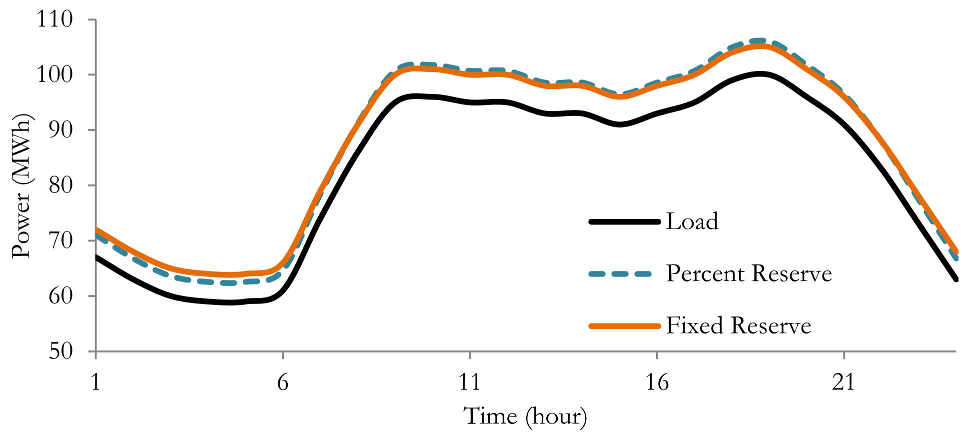

Figure 3-1 Examples two types of reserve: fixed percentage of load and fixed MW

These issues are all important to assess, and it is not often that a system operator can implement many changes at once. This study evaluates specific improvements to reserve procurement, comparing which have the greatest impact on market efficiency. The improvements directly or indirectly involve each of the three stages of operating reserves procurement, and compare an existing method with a proposed method. In the Netherlands, the reserve requirement is presently based on a seasonal average, and capacity is acquired by the TSO months in advance through long-term contracts with generators within the Netherlands alone [89].

In the alternative reserve schemes proposed in this chapter, three enhancements are considered. The first improvement changes the quantity of reserve procured from a seasonal fixed amount to a value that changes day-by-day. In the summer months, the seasonal quantity of reserve procured might be high due to expected high temperature during peak summer weeks.

[17] Deliverability of a flexible ramping product is currently an issue in California in the Energy Imbalance Market, the region's real-time market [235].

However, temperatures in early weeks might be lower than expected, causing over-procurement of reserve. The suggested improvement would procure reserve based on the following day's expected load; this should provide a tighter bound compared to a seasonal average. The second improvement is changing the allocation of reserve from long-term contracts to a market mechanism. The suggestion creates a system similar to the U.S., which co-optimizes energy and reserve together. The allocation of reserve can follow the least cost dispatch and take advantage of online generation, rather than contracting with one particular generator. The final improvement coordinates the allocation and activation of reserve. Each country allocates and activates reserve independently, with minimal real-time balancing. This improvement would combine the reserve needs of neighboring countries to take advantage of less expensive generation. The three enhancements are tested in an hourly market model simulation for Europe for a year of wind data given the expected penetration for the year 2030. Background on reserve products and European markets can be found in Section 3.2, followed by a detailed description of the three improvements and mathematical formulation used in Section 3.3. The simulation data and results can be found in Section 3.4 with a discussion of the implications and limitations in Section 3.5.

3.2 BACKGROUND

There are many types of operating reserve depending on their use, and many conflicting terms across system operators. Reserve types can be distinguished by the response time, physical characteristics, and the type of event to which it is responding [93]. Most European markets have primary, secondary, and tertiary reserve [94]. Primary reserve, also called governor response and regulating reserve, responds within seconds or less and provide frequency support. After an event occurs, there is usually slow tertiary reserve which replaces reserve that was used for a contingency. Between these types is secondary reserve, a type of reserve that responds within seconds to minutes to ease forecast uncertainty and in some cases, follow load. These types of reserve also fall into categories of spinning and non-spinning reserve, where spinning reserve is provided by an online generator and non-spin by a resource that can come online quickly. Most U.S. markets have further distinctions between types of reserve. Contingency reserve is only used in an emergency setting and load following reserve is used in 'non-event' situations, generally for balancing due to forecast errors. Generally, all regulation reserve or primary reserve must be spinning, while contingency reserve might be spinning or non-spinning. The distinctions between countries and markets is wide ranging; several comparison tables can be found in [88], [93], [95]. In the Netherlands, secondary reserve is called automatic or manual frequency response (aFRR or mFRR) [96],[97], the latter of which is the focus of reserve procurement in this chapter. This section will first detail broad literature on reserve modeling, particularly the choice of reserve requirement. Next, specific literature on European institutions, markets, and procedures will be explained, including studies in the literature and from collaborators most similar to the work in this chapter.

Reserve modeling has received increased attention in the literature due to the uncertainty from wind and solar generation. There has been a great deal of literature discussing alternative reserve models that directly consider uncertainty. A review of reserve markets including a focus on both modeling issues and technical constraints can be found in [94]. Literature on modeling extends from competitive market models to theoretical equilibrium models. More recently, a complementary review of flexibility was published including flexibility metrics, market design, and the possibility of distribution system operator (DSO) interactions [98]. Finally, a comprehensive review of different types of reserves in a high wind system can be found in [93], with country-specific practices and policies and an evaluation of reserve in wind integration studies.

Much of the literature on reserve proposes new methodologies for determining the reserve requirement needed [99]–[103]. The authors of [99] propose a new reserve requirement based on the loss of load costs associated with each period, with only slightly increased computational time. The method is generalized in [100], where demand is represented as a probability distribution and Monte Carlo analysis is used to compare the proposal against traditional reserve formulations. Another proposed method in [101] for reserve requirement uses a probabilistic approach that produces an hourly requirement that reduces the risk of load shedding over the year. While also incorporating the cost of lost load, the authors in [102] explicitly consider probabilities with a stochastic two stage model, acknowledging the computation burden is high and proposing methods to decrease the difficulty. In [103], probabilistic sizing is used to determine reserves that directly considers the cost of activation. Finally, rather than a new proposal, [104] compares two common requirements, N-1 or loss of a single generator or network device, and 3σ or three times the standard deviation for demand and renewables. With a common test system, the authors find the amount of reserve schedule depends on the penetration of wind, with levels decreasing as penetration increase until they reach a minimum and finally increase with penetration. They point to the utilization of a large nuclear plant, which can impact total reserve need in a small system. The majority of these methods use a benchmark system, one of the IEEE Reliability Test Systems, to analyze proposed reserve procurement systems [99], [100], [102]. This chapter distinguishes itself by utilizing a real-world system and historical data instead of a well-known test case, focusing on European markets. Each paper cited proposes a new method for determining the requirement that improves economic efficiency or increases wind penetration. The focus of this chapter is comparing improvements beyond a single requirement (endogenous to the day-ahead market model). These methods can be used as future comparisons to the simple requirement used in this chapter, which is one of the three suggested improvements.

The literature on European reserve or ancillary services markets is extensive. For those more familiar with U.S. markets, [105] offers a comparison between U.S. and European markets, with a summary figure of European markets and the role of the TSO. An overview of the European ancillary services market can be found in [106], details of generator decision making in the

Nordic market detailing reserve processes in [107], an overview of Dutch markets can be found in [108] and [109], and a Dutch a wind integration study in [110]. In [108], they show there has been more over-procurement than under-purchasing of reserve in a twelve year period,[18] suggesting a tighter bound could lead to efficiency improvements. This claim helps motivate the reserve requirement improvement, showing that rather than use a broad requirement, a tighter requirement might benefit the system (described in Section 3.3.2.c).

In analyzing the future European electricity system, [111] promotes the change from bilateral contracts for reserve to a market-based system, one of the improvement suggested in this chapter (Section 3.3.2.b).[19] The authors in [112] suggest design options for such a market, distinguishing between longer-term reserve capacity and reserve needed for balancing; similar to how this chapter refers to the different phases of reserve: allocation and activation. They delve further into the details of implementing a reserve market, where the Common Merit Order method is most similar to the proposal in Section 3.3.2.b. The basis for [113] also assumes the existence of a simultaneous energy and reserve market, and additionally examines coordination, similar to Section 3.3.2.d.

Several papers emphasize the need for cross-border integration and interconnection as an essential tenant of renewable integration in Europe [111], [113], [114]. Those authors' assertion complements the last improvement suggested in this chapter, the coordination of reserve across countries (Section 3.3.2.d). The framework used in [113] is most similar to the improvements suggested in this chapter, coordinating reserve across Northern Europe and modeling a common marketplace. With both system-wide reserve allocation and activation across the region in balancing markets, reserves from the Nordic region are highly utilized in the remainder of northern Europe, where 30% of the requirement is traded across the border between the Nordic and UCTE systems which are not synchronized and so are connected with DC lines. In a follow-

[18] Under-procurement might also lead to loss of load, suggesting proper penalties and incentives are needed to ensure the system is long rather than short on reserve.
[19] Co-optimization of energy and reserve is current practice in all U.S. markets.

up study, [114] shows that allocating transmission capacity is a vital aspect of cross-border reserve trading.

The topic is also explored through my collaborative work with a visiting student to Johns Hopkins, where the resulting paper, [92], examined the coordination of reserves for four countries. The simulations found that the coordinated allocation of reserve in the day-ahead market did not necessarily correspond to lower system costs in the balancing market. Without reserving capacity on transmission lines, the reserved power was not always deliverable when needed. This result leads to the addition of extra simulations to confirm the outcome using a European-wide model.

There are similarities between this chapter and the papers in the previous paragraphs, as all examine coordination between regions in Europe and use of a market setting to procure reserve. The modeling frameworks in [79], [100], and [101] are most similar to the methodology in this chapter. All use a unit commitment model to simulate the day-ahead market, followed by an activation of reserve in a balancing market. The input data and networks that are utilized differ, but the basic framework is the same. This work is distinguished by analyzing the combination of these proposals in addition to alternate reserve requirements or coordination alone. By comparing three modes of improvements, this chapter can assess which improvement makes the biggest impact for maximizing market surplus and integrating wind power. While there is literature on specific improvements to one aspect of reserve modeling, no paper compares improvements to different steps of the reserve process in addition to contract versus market-based reserved. This project offers a comparison of reserve improvements in procurement allocation, and activation focusing on the Dutch and European markets. While the case study is specific to a region, the method may be used to study other systems and countries.

3.3 METHODOLOGY

In order to deal with the uncertainty that rises from wind generation, a model is needed that represents the multi-stage nature of operational decisions and the uncertainty and variability at each stage. The Netherlands does not exist in isolation, and therefore the model must be able to interact with its neighbors and reflect current market conditions. For these reasons, we chose to run simulations for this project with the Comprehensive Market Power in Electricity Transmission and Energy Simulator (COMPETES) expanded to include a day-ahead unit commitment stage [115], [116]. COMPETES is a network-constrained pan-European market model with one node per country. As seen in Figure 3-2, the nodes consist of 26 European Union member states and 7 non-EU countries. The full list of countries and the abbreviations used in the results section can be found in Table 3-1. The model was originally created as a game-theoretic model of generation dispatch on a power network by researchers from Johns Hopkins and Energy Research Center of the Netherlands (ECN) [117], [118]. COMPETES includes a combination of mixed-integer and relaxed unit commitment formulations for day-ahead unit commitment, along with operating reserve requirements, transmission constraints, and rules for thermal generation in both day-ahead and balancing (real-time) markets.

Section 3.3.1 describes the modeling framework for the three improvements. The formulation for the unit commitment model is presented in detail in Section 3.3.2, with subsections for each type of improvement. Section 3.3.3 explains the sensitivity performed considering the renewable forecast error.

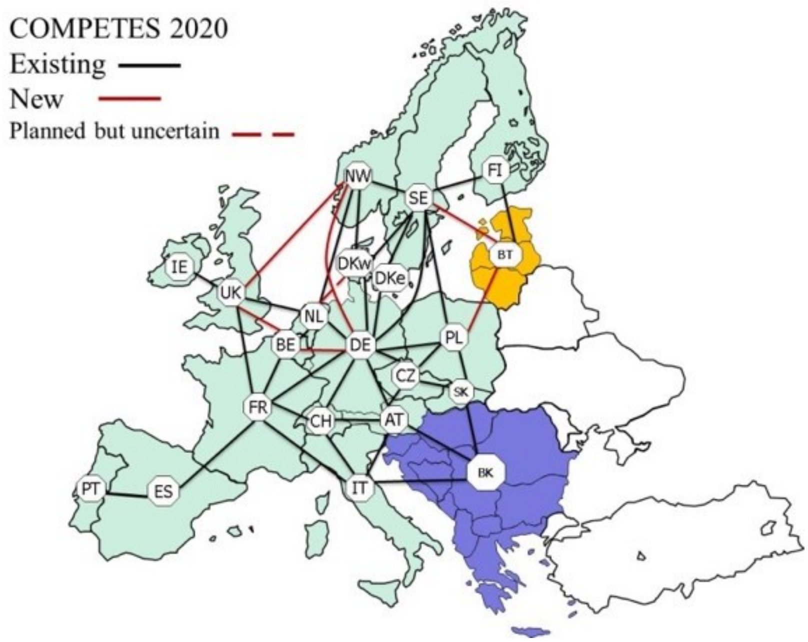

Figure 3-2 Network in COMPETES model

Table 3-1 Countries in COMPETES and their abbreviations

BEL	Belgium	POR	Portugal
CZE	Czech Republic	SKO	Slovakia
DEN	Denmark East	SPA	Spain
DEW	Denmark West	SWE	Sweden
FIN	Finland	UKI	United Kington
FRA	France	SWI	Switzerland
GER	Germany	NOR	Norway
IRE	Ireland	BLK	Balkans
ITA	Italy	BLT	Baltics
NED	Netherlands	AUS	Austria
POL	Poland		

3.3.1 Modeling Framework

There are three improvements to reserve procurement that will be tested. Each improvement roughly captures a different step in the reserve process and contrasts current practice with what we hypothesize to be an improvement. The first proposes to update the allocation phase from using bilateral contracts to a market system (referred to as 'type'). The second is an improvement to the sizing or procurement of reserve (referred to as 'reserve requirement'). The third improvement coordinates reserve allocation and activation among countries in northwest Europe (referred to as 'coordination'). This framework is unique among the literature. There have been papers that compared one or two of these improvements, but to my knowledge, this combination and extensive simulation has not been performed previously.

Figure 2 shows a comparison of different simulations with increasing complexity moving away from the origin. The costliest point is hypothesized to be the black dot, representing a limited set of generators provided by contract for a seasonal requirement. The least costly simulation is hypothesized be the market-based daily reserve requirement including all generators, represented by a star. Moving away from the origin, the complexity increases. Determining reserve on a daily basis takes more computational effort than a seasonal requirement. Similarly, operating a market is more intensive than bilateral contracts. While complexity increases, I also hypothesize efficiency increases. The improvements should increase market surplus through the different simulations.

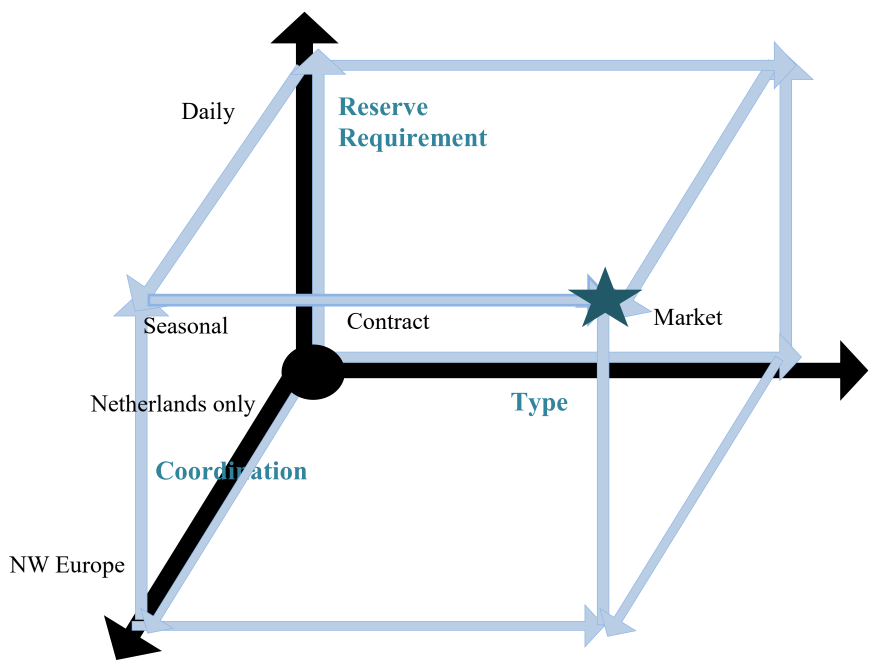

Figure 3-3 Conceptual diagram of reserve improvements, where complexity and efficiency increase moving away from the origin. The type improvement is described in 3.3.2.a, the reserve requirement is in 3.3.2.b, and the coordination is in 3.3.2.c.

The day-ahead and balancing models run sequentially to mimic actual operations. The day-ahead model fixes the generation commitment schedule, then the real-time model runs hour by hour with that schedule. Generators that are able to start up quickly will be able to commit in the balancing market, while slow generation commitment will either be fixed based on the day-ahead schedule or out of the market. The day-ahead market uses wind forecasts scaled based on historical data and the real-time market uses actual data where available and similarly distributed data where not available. A further description of the data used for wind can be found in 3.3.3.

Although there are twelve possible simulation combinations from Figure 3-3, all are not practical. Reserve is currently contracted on a seasonal basis; it is highly unlikely that contracts will change daily to accommodate a new daily reserve requirement. Therefore, the daily contract-based simulations (with and without coordination) were not simulated. Additionally, the seasonal coordinated day-ahead simulation was eliminated as it would require companies across countries to coordinate contracts. Although possible, it would require many assumptions about cross-country cooperation. In total, the eight simulations done can be found in Table 3-2.

Table 3-2 Eight simulations performed using COMPETES

		No Coordination	Balancing Only	DA and Balancing
Contract	Seasonal	X	X	
	Daily			
Market	Seasonal	X	X	X
	Daily	X	X	X

3.3.2 Model Formulation for Day-Ahead and Real-Time Markets

Each of the eight scenarios is simulated for both a day-ahead and balancing market. The formulation for the day-ahead market is described below, followed by a description of the differences between the day-ahead and balancing constraints. The specific differences between the scenarios are described in subsections 3.3.2.a, b, and c. The definitions of sets, variables, and parameters can be found in the Nomenclature section at the beginning of the dissertation and details about the input parameters can be found in Section 3.4.

Equations (3-1)-(3-21) are considered the basic day-ahead unit commitment model for this project and will be used as the base model for all scenarios. Among the simulation types, the formulation shows a co-optimized energy and reserve market; differentiated constraints for additional simulations are shown after the formulation. The objective of the model is to minimize operating costs, which consist of fixed costs for start-up and minimum-load operation, as well as the marginal cost for the power dispatched. The objective in (3-1) is straightforward; minimize operating costs, which include the cost of generation, storage and lost load. The value of lost load (VOLL) is €3000/MWh, which is the market cap for most European markets [119] and the cost of storage is based on the investment costs [118]. Constraints (3-2) and (3-3) limit the power capacity of each generator. Due to the large problem size, the unit commitment formulation is based on [120] for the resources in the Netherlands. This formulation for unit commitment produces the same result as other unit commitment formulations, but is shown to be tighter and more compact. The generation dispatch variable is constrained between 0 and the maximum less the minimum capacity. Many other unit commitment formulations limit the dispatch between minimum and maximum ($P_g^{\min} u_{g,t} \le p_{g,t} \le u_{g,t} P_g^{\max}$).

For resources outside the Netherlands, the generation is aggregated by year and fuel type, meaning there are not individual generators. Committing the aggregated resources using a strict binary variable would imply that the all generators of a particular type would be committed, where this is not the case in reality. It would also be a significant computational burden. Therefore, for all generators outside the Netherlands, the same formulation is used except the unit commitment variables are relaxed between 0 and 1. This allows generators that would otherwise be operating at their minimums to set price, a characteristic that provides further insight into pricing and echoes the ELMP pricing model from Section 2.6.1.

The imports and exports between countries are described in constraints (3-4)–(3-6), with Net Transfer Capacity (NTC) limits on lines and export/import limits between countries. A transshipment ("pipes-and-bubbles") formulation is used. Constraint (3-7) defines the startup and commitment status of the generator, where the variables *u, v,* and *w* can either be 1 (startup/commitment occurs in that interval/shutdown) or 0 (otherwise).[20] The ramping capability of the generators is defined in (3-8) and (3-9), where (3-8) limits ramping up and (3-9) limits ramping down. The reserves are found in the ramp up but not ramp down constraint. While it can be added in both directions, the characteristics of the system do not need additional ramp down from reserve, since ramp up is the greater concern. Constraints (3-10) and (3-11) define the minimum up and down times for the generators during the 24-hour commitment period. Similar to the dispatch model, the wind injection can be curtailed in the node balance constraint (3-12), and is limited by the day-ahead forecast in (3-13). Demand can also be shed at a high cost, and the slack variable's upper bound is shown in (3-14).

The constraints describing energy storage in the network are in (3-15)–(3-18). The charging and discharging is limited by the maximum power output of the storage resource in (3-15) and (3-16). The state of charge variable for storage has an upper bound of the maximum capacity in (3-17). The storage balance constraint, which determines the state of charge between periods relative to the charging and discharging, is found in (3-18). Reserve for storage is similar to the formulation in [121], with the reserve variable limited by the amount discharged and the production capacity in (3-16), and by the state of charge in (3-19). There are other formulations for storage reserve [122], many of which are related to electric vehicle charging [123], [124]. Some formulations assume reserve must be utilized to a certain degree in each period [122];

[20] The shutdown variable *w* does not have a cost in the objective, and both *v* and *w* could be continuous between 0 and 1. However, the formulation in [120] proves that if both are binary the formulation is tight and compact.

however, the formulation presented below might overestimate the contribution of storage to reserves. Future work in this area would simulate several formulations to determine how much storage is overestimated and how that would impact operating costs.

Constraint (3-20) defines the requirement for operating reserves, $R_{i,t}^{req}$. The requirement is dependent on which of the two scenarios is being simulated. The definition of $R_{i,t}^{req}$ is found in Section 3.3.2.c. The value of requirement is based on the NREL 3+5 rule [88], where 3% of demand and 5% of variable generation is required. The rule was originally intended for contingency reserve, with half available as spinning reserve. Since this model only considers spinning reserve and has a high renewable penetration, the entire 3% and 5% are the requirement utilized. A report from IEA determined that 4% of installed wind would be necessary for additional reserves with penetrations under 10% [9]. Since the penetration in most countries exceeds 10%, a requirement of 5% is not unreasonable. Finally, the commitment variable defining the status of a generator is restricted to be a binary variable in (3-21) for Dutch generators and relaxed for remaining generators in (3-22). The variables that are nonnegative are designated in (3-23). All notation definitions can be found in the Notation section at the beginning of the dissertation. For any one simulation, the time horizon is twenty four hours, with the last hour of the previous day used as input for the next.

Day-Ahead Unit Commitment in COMPETES

$$\min \sum_{t} \sum_{g} \left(c_g p_{g,t} + c_g^{\mathrm{SU}} v_{g,t} + c_g^{\mathrm{NL}} u_{g,t} \right) + \sum_{i} VOLL\left(d_{i,t}^{\mathrm{curt}}\right) + \sum_{v} c_v^{st} \left(\frac{s_{v,t}^{\mathrm{dc}}}{\eta_v} \right) \tag{3-1}$$

Subject to

$$p_{g,t} + r_{g,t}^{\mathrm{SP}} \le u_{g,t}\left(P_g^{\max} - P_g^{\min}\right) - v_{g,t}\left(P_g^{\max} - R_g^{SU}\right) - w_{g,t+1}\left(P_g^{\max} - R_g^{SD}\right), \ \forall g, t \tag{3-2}$$

$$p_{g,t} \ge 0, \ \forall\, g, t \tag{3-3}$$

$$NTC_k^{\min} \le f_{k,t}^{\mathrm{up}} - f_{k,t}^{\mathrm{dn}} \le NTC_k^{\max}, \ \forall\, k, t \tag{3-4}$$

$$Import_i^{\max} \le \textstyle\sum_{\forall k} I_{i,k}\left(-f_{k,t}^{\mathrm{dn}} + f_{k,t}^{\mathrm{up}}\right), \ \forall\, i, t \tag{3-5}$$

$$\sum_{\forall k} I_{i,k}\left(-f_{k,t}^{\mathrm{dn}} + f_{k,t}^{\mathrm{up}}\right) \le Export_i^{\max},\ \forall\, i,t \tag{3-6}$$

$$v_{g,t} - w_{g,t} = u_{g,t} - u_{g,t-1},\ \forall\, g,t \tag{3-7}$$

$$p_{g,t} + r_{g,t}^{\mathrm{SP}} - p_{g,t-1} \le R_g,\ \forall\, g, t > 1 \tag{3-8}$$

$$p_{g,t-1} - p_{g,t} \le R_g,\ \forall\, g, t > 1 \tag{3-9}$$

$$\sum_{r=t-\tau_g^{\mathrm{UT}}+1}^{t} v_{g,r} \le u_{g,t},\ \forall\, g, t \ge \tau_g^{\mathrm{UT}} \tag{3-10}$$

$$\sum_{r=t+1}^{t+\tau_g^{\mathrm{DT}}} v_{g,r} \le 1 - u_{g,t},\ \forall\, g, t \le |T| - \tau_g^{\mathrm{DT}} \tag{3-11}$$

$$\begin{aligned}\sum_{k} I_{i,k}\left(f_{k,t}^{\mathrm{dn}} - f_{k,t}^{\mathrm{up}}\right) + \sum_{g\in G_i}\left(p_{g,t} + u_{g,t}P_g^{\min}\right) + w_{i,t}^{\mathrm{inj}} + \sum\nolimits_{v\in V_i} s_{v,t}^{\mathrm{dc}} - s_{v,t}^{\mathrm{ch}} \\ = D_{i,t}^{\max} - d_{i,t}^{\mathrm{curt}},\ \forall i,t\end{aligned} \tag{3-12}$$

$$0 \le w_{i,t}^{\mathrm{inj}} \le W_{i,t},\ \forall\, i,t \tag{3-13}$$

$$0 \le D_{i,t}^{\max} \le D_{i,t}^{\max},\ \forall\, i,t \tag{3-14}$$

$$s_{v,t}^{\mathrm{ch}} \le C_{v,t}^{\max},\ \forall\, v,t \tag{3-15}$$

$$s_{v,t}^{\mathrm{dc}} + r_{v,t} \le C_{v,t}^{\max},\ \forall\, v,t \tag{3-16}$$

$$0 \le l_{v,t} \le S_{v,t}^{\max},\ \forall\, v,t \tag{3-17}$$

$$l_{v,t} - l_{v,t-1} - \eta_v s_{v,t}^{\mathrm{ch}} + \frac{s_{v,t}^{\mathrm{dc}}}{\eta_v} = 0, \qquad \forall\, v,t \tag{3-18}$$

$$\frac{r_{v,t}}{\eta_v} \le l_{v,t}, \qquad \forall v,t \tag{3-19}$$

$$\sum_{g\in G_i} r_{g,t}^{\mathrm{SP}} + \sum_{v\in V_i} r_{v,t} \ge R_{i,t}^{req},\ \forall i,t \tag{3-20}$$

$$u_{g,t}, v_{g,t}, w_{g,t} \in \{0,1\},\ \forall i \in G^{NL}, t \tag{3-21}$$

$$0 \le u_{g,t}, v_{g,t}, w_{g,t} \le 1,\ \forall i \in G^{EU}, t \tag{3-22}$$

$$f_{k,t}^{\mathrm{up}}, f_{k,t}^{\mathrm{dn}}, s_{v,t}^{\mathrm{dc}}, s_{v,t}^{\mathrm{dc}}, r_{v,t}, l_{v,t} \ge 0, \forall v,k,t \tag{3-23}$$

The Europe-wide balancing market model is based on the day-ahead formulation with adjustments for reserve and fast generators. This chapter uses an hourly balancing market model, mimicking the last intra-day market, which is typically operated one-hour before delivery [117].

In this simulation, reserves are released, meaning all reserve variables and constraints (3-19) and (3-20) are not included. The formulation for the day-ahead market is used for fast-start resources, or those generators that have a minimum run time of one hour and can ramp to their minimum capacity within an hour. In actual market operations, these resources may be required to start up within a shorter time period, however this balancing market is modeled hourly. The slow generation resources have fixed commitment status, meaning the $u_{g,t}$, $v_{g,t}$, and $w_{g,t}$ variables are held at their day-ahead status. The dispatch level for the units can change, still limited by their ramping capability. In some of the simulations, cross border flows are fixed, while in others the set of resources that were formerly contracted for reserve are available for dispatch. These distinctions can be found in the following three subsections, one for each improvement. Figure 3-3 shows an experimental design in which each of three changes are considered and varied between two options (one the base case, and the other an improvement). In the next three subsections I explain each of these characteristics.

3.3.2.b Reserve Characteristic 1: Contract- vs. Market-Based Procurement

The first of three proposed improvements considered in this chapter is the method by which reserves are allocated, where current practice allocates based on long-term contracts with particular generators. The improvement to current practice is allocation of reserve through an auction, where energy and reserve are co-optimized in the day-ahead market (as is done in U.S. markets). The hypothesis is the market-based case will provide a lower cost solution, since it will allocate reserve closer to delivery (with more information on the state of the system) and based on cost rather than contracts.

Reserve contracts are usually between the TSO and a company or generation owner. As the details of contracts are private information, the modeling assumes generators engaged in a contract will be paid a fixed price per MWh for the duration of the contract [89], [108]. The quantity of reserves contracted is the same as the reserve requirement in the market-based model. Before any simulation is done, a simple algorithm determines which generators will be contracted for reserve based on their installation date and fuel type. For the simulations, only fossil fuel generators were contracted. The naïve algorithm first contracted the oldest oil, coal, gas, and lignite, followed by newer units until the requirement was met. A unit was contracted for its full capacity; if the requirement was exceeded due to the last unit added, it could not be larger than 110% of the requirement. For example, if the requirement was 1000 MW and 950 MW had been allocated, a generator larger than 150 MW would be skipped for a newer resource less than 150 MW. All countries procured resources in this manner except Switzerland, which had not fossil fuels. In that case their hydro resources were partially contracted to provide reserve.

The generators in each country that are contracted for reserve would be excluded from participation in the day-ahead market, since their capacity could not be scheduled for energy generation. In the balancing market, the reserves are released, and can be used for balancing the wind forecast error. Other than fixing the off status of the contracted generators, no other modeling is altered from the formulation above. The contracted generators would be paid a price for their capacity as reserve, and the balancing price for any energy provided in that market. Since information on contract prices is not publicly available and would greatly impact cost calculations for comparison, no reserve costs were included in the results. A further discussion can be found in Section 3.4.5.

3.3.2.c Reserve Characteristic 2: Seasonal vs. Daily Reserve Requirement

The second improvement analyzed is changing the quantity of reserve requirement itself. Current practice sets a seasonal reserve quantity, and the proposed improvement would set a daily requirement. The hypothesis is the daily requirement would produce a lower cost solution compared to the seasonal requirement, since the amount would be tailored to the next day's needs rather than a seasonal estimation. In both the seasonal and daily case, the requirement utilized is simple in order to compare the importance of relative size of the requirement and ensure ease of implementation. As mentioned in the formulation section, the NREL 3+5 rule for load and renewable energy is used for both seasonal and daily requirements [88]. The requirement was written for use in a U.S. market, where contingency reserve is procured using both spinning and non-spinning reserve. In this modeling framework, only spinning reserve (reserve that is online) is procured. The rule is adapted to set 3% of average load and 5% of average renewable forecast is the minimum requirement. These values can be updated in future simulations to reflect wider or narrower ranges. A rule using 1% of load and 3% of renewables was tested, and the resulting load shed in the simulation was greater than using 3% and 5%.

The seasonal simulation is based on the average load and renewable forecast for winter, spring, summer, and fall, leaving 4 different requirements for the year. The daily requirement is based on the average load and wind forecast for the following day, leaving 365 different requirements for the year. The reserve requirement is determined exogenously from the model, since it is an input to the day-ahead market. Both requirements are formulated below.

Seasonal Requirement

$$R_{i,s}^{req} = \frac{0.03 \sum_{\forall t_s} d_{i,t_s}^{forecast} + 0.05 \sum_{\forall t_s} W_{i,t_s}^{forecast}}{T^{season}}, \qquad \forall i, s \in \{winter, spring, summer, fall\}$$

where T^{season} is the number of hours in either winter, spring, summer or fall, and t_s is the set of hours in each season *s*.

Daily Requirement

$$R_{i,d}^{req} = \frac{0.03 \sum_{\forall t_d} d_{i,t_d}^{forecast} + 0.05 \sum_{\forall t_d} W_{i,t_d}^{forecast}}{T^{day}}, \qquad \forall i, d \in \{1,2, \dots ,365\}$$

where T^{day} is the number of hours in in the day (24), and t_s is the set of hours in each day *d*.

3.3.2.d Reserve Characteristic 3: Independent Allocation vs. Coordination among European Countries

The third improvement deals with the coordination of reserve procurement between the Netherlands and its neighbors. In current practice, each country procures reserve from resources within its boundaries [89]. With some coordination between neighbors, the reserve can be procured from the least cost resource, possibly leading to higher overall system efficiency. The hypothesis is the coordinated scenarios will produce a lower cost result and integrate more wind energy. The overall system should be at least the same or better off with coordination in the balancing market, since the solution without coordination is part of the set of solutions in the coordinated case. This type of coordination is also explored in Chapter 5 for energy markets. However, coordination across both day-ahead and balancing markets leads to more complex interactions.

For the market-based simulations, all reserves will be allocated to specific generators in the day-ahead market and released in the balancing market, being allowed to freely provide energy according to its energy cost. Actual markets would hold reserve in the balancing market for use in real-time delivery; however, this chapter treats the balancing market as activation in real-time. The first reason for this modeling assumption is availability of data. Simulation of real-time delivery would require additional wind and load data based on an hour-ahead forecast. This data was not available for countries in this study. Second, real-time simulations would best be served using an ac model so voltage and reactive power deviations could be captured. The model used for this chapter is meant to analyze market mechanisms and not real-time operations. For this reason, reserves meant to manage second-by-second frequency variations (primary reserves) are also not modeled.[21]

The no coordination simulation will allocate country by country, while the improvement will combine Northwest Europe, expanding the reserve zone to the countries surrounding the Netherlands. The balancing market will then incorporate those countries, while fixing any imports and exports. The countries coordinating in the simulation are sometimes referred to as Central Western Europe, and were the focus of previous work [92]; they include the

[21] Unlike the U.S., there are no contingency or ramping reserve products in ENSTO-E. The response to a contingency is provided by primary, secondary and tertiary reserves, which will respond at different speeds [88].

Netherlands, Belgium, France, and Germany. Due to increased trade,[22] the simulations also include the United Kingdom; future studies can analyze different combinations of these countries to evaluate if any supersede the five used in simulations.

Previous research in [92] identified an important discrepancy between expected benefits from coordination in the day-ahead and balancing markets, to coordination in the balancing market alone. In [92], the simulations where coordination occurred in both allocation phases (day-ahead market) and the activation phase (balancing market) resulted in higher costs compared to coordination in the activation phase alone. To investigate this discrepancy in COMPETES, an additional set of scenarios is created to account for coordination in the balancing market alone. This improvement was updated to have three levels of coordination: (1) no coordination, where each country allocates and activates reserve independently, (2) coordination in the balancing market, where allocation of reserves in the day-ahead market occurs independently by country, and (3) coordination in day-ahead allocation and balancing market activation.

For the case with no coordination, each country procures reserve independently and balances forecast error independently. Constraint (3-20) is modeled for each country *i* and the balancing market fixes the total imports and exports between countries. When only the balancing market is coordinated, the day-ahead market remains independent, as in the no coordination scenario. Then in the balancing market, Northwestern European countries can trade between each other, while remaining countries balance independently and their imports and exports are fixed. For full coordination, the day-ahead market is also coordinated; any generator in Northwestern Europe can contribute to the reserve requirement following NTC limitations. The remaining countries have separate reserve requirements that must be fulfilled with in-country resources. The balancing market allows trade among Northwest Europe, again fixing the imports

[22] Trade between the UK and the Netherlands increased due to the BritNed line, an underwater high voltage dc transmission line that began operation in 2011 [236].

and exports of all other countries. The modifications are shown below, where N^{NWE}, K^{NWE} is the combined set of countries or the set of lines connected the countries in Northwest Europe: the Netherlands, Belgium, Germany, France, and the United Kingdom. The set $-N^{NWE}, -K^{NWE}$ contains all other countries and the inter-tie lines connecting them, see Table 3-1 for a complete list.

No Coordination

- **Day-Ahead**: reserve constraint for each country, same as (3-20)

  All countries schedule independently

  $$\sum_{g \in G_i} r_{g,t}^{\mathrm{SP}} + \sum_{v \in V_i} r_{v,t} \geq R_{i,t}^{req}, \ \ \forall i \in N, t$$

- **Balancing**: eliminates constraints (3-4)–(3-6) and fixes inter-tie exchanges, $f_{k,t}^{\mathrm{up}}, f_{k,t}^{\mathrm{dn}}$

  All countries balance independently

Balancing-Only Coordination

- **Day-Ahead**: reserve constraint for each country, same as (3-20)

  All countries schedule independently.

  $$\sum_{g \in G_i} r_{g,t}^{\mathrm{SP}} + \sum_{v \in V_i} r_{v,t} \geq R_{i,t}^{req}, \ \ \forall i \in N, t$$

- **Balancing**: modifies constraints (3-4)–(3-6), fixes inter-tie flow, $f_{k,t}^{\mathrm{up}}, f_{k,t}^{\mathrm{dn}}\ \forall k \in K^{-NWE}$

  Counties outside of Northwest Europe balance independently, countries in Northwest Europe balancing together (inter-tie lines are not fixed)

  $$NTC_k^{\min} \leq f_{k,t}^{\mathrm{up}} - f_{k,t}^{\mathrm{dn}} \leq NTC_k^{\max}, \ \forall\, k \in K^{NWE}, t$$

  $$Import_i^{\max} \leq \sum_{\forall k \in K^{NWE}} I_{i,k}\left(-f_{k,t}^{\mathrm{dn}} + f_{k,t}^{\mathrm{up}}\right), \ \forall\, i \in N^{NWE}, t$$

  $$\sum_{\forall k k \in K^{NWE}} I_{i,k}\left(-f_{k,t}^{\mathrm{dn}} + f_{k,t}^{\mathrm{up}}\right) \leq Export_i^{\max}, \ \forall\, i \in N^{NWE}, t$$

Day-Ahead and Balancing Coordination

- **Day-Ahead**: single reserve constraint for Northwest European region and independent constraints for all remaining countries, modifies (3-20)

Counties outside of Northwest Europe schedule reserves independently, countries in Northwest Europe schedule reserves together through a single constraint

$$\sum_{g \in G_i} r_{g,t}^{\mathrm{SP}} + \sum_{v \in V_i} r_{v,t} \geq \sum_{i=N^{NWE}} R_{i,t}^{req}, \quad \forall i = N^{NWE}, t$$

$$\sum_{g \in G_i} r_{g,t}^{\mathrm{SP}} + \sum_{v \in V_i} r_{v,t} \geq R_{i,t}^{req}, \quad \forall i \in N^{-NWE}, t$$

- **Balancing**: modifies constraints (3-4)–(3-6)

  Counties outside of Northwest Europe balance independently, countries in Northwest Europe balancing together (inter-tie lines are not fixed)

$$NTC_k^{\min} \leq f_{k,t}^{\mathrm{up}} - f_{k,t}^{\mathrm{dn}} \leq NTC_k^{\max}, \ \forall\, k \in K^{NWE}, t$$

$$Import_i^{\max} \leq \sum_{\forall k \in K^{NWE}} I_{i,k}\left(-f_{k,t}^{\mathrm{dn}} + f_{k,t}^{\mathrm{up}}\right), \ \forall\, i \in N^{NWE}, t$$

$$\sum_{\forall kk \in K^{NWE}} I_{i,k}\left(-f_{k,t}^{\mathrm{dn}} + f_{k,t}^{\mathrm{up}}\right) \leq Export_i^{\max}, \ \forall\, i \in N^{NWE}, t$$

3.3.3 Role of Forecast Error between Day-Ahead and Balancing

In addition to the three improvements described in the previous three subsections, scenarios for the actual realization of wind for real-time markets are developed to ensure a comprehensive set of results. There is uncertainty between the day-ahead and balancing markets. The midday forecasts are between 12 and 36 hours from delivery considering the auction runs for the following day (midnight to midnight) [125]. Due to the error from forecasts, simulation of a single actualization might not capture the full uncertainty of wind. The simulation of multiple actualizations ensures the results reflect more than a single realized uncertainty. The amount of reserve procured in the day-ahead market might be enough for one actualization but might fail in other cases. Analyzing the average of five real-time actualizations helps capture that uncertainty.

As discussed earlier, only wind input data changes between the day-ahead and balancing markets, load data remains constant. This study focuses on the impacts of wind, and the simulations highlight the change of wind forecast error compared to realizations, meaning one uncertainty is simulated to isolate the impact. The data available is also limited for load, and synthesizing data for actualizations would require many assumptions. If further data becomes available, future studies can investigate the effect of interactions between wind and load forecast and actualizations [126], [127].

The balancing wind data is based on historical data with additional scenarios created using an autoregressive (AR) model [128]. The production data for actual realizations is provided by the Wind Unit at ECN from public and private wind data, and the AR model was developed by a colleague at ECN, Özge Özdemir. The available historical data is used for the first of five balancing market wind realizations. The wind forecasts and actualizations are then used to create four additional synthetic annual wind time series for use as scenarios in the balancing market. The forecast less the wind actualization calculates basic forecast error, which is used in an autoregressive model. The first autocorrelation, ρ, is calculated in (3-24), where x_t is the forecast error, or the forecast less the actual realization, μ is the mean forecast error, σ is the standard deviation, and T is the total number of time periods, or 8760 hours. The noise, ϵ, is calculated using a normal distribution with mean zero and standard deviation $\sigma\sqrt{1-\rho^2}$, shown in (3-25). Finally, the scenarios are calculated using an AR(1) model in (3-26), where the parameters are the series autocorrelation and mean. For countries without available data, the average mean and autocorrelation is used to calculate noise and the additional scenarios.

$$\rho = \frac{\frac{1}{T}\sum_t (x_t - \mu)(x_{t-1} - \mu)}{\sigma^2} \tag{3-24}$$

$$\epsilon_t = Norm\left(0, \sigma\sqrt{1-\rho^2}\right) \tag{3-25}$$

$$x_t^S = \mu + \rho(x_{t-1} - \mu) + \epsilon_t \tag{3-26}$$

As shown in Figure 3-4, one wind forecast is used for the day-ahead market. The simulation results for scheduling and inter-tie line flow are fed into five separate annual wind actualizations, each 365 days, for the balancing market. Five different yearly simulations run one after the other, each using the same day-ahead input, meaning for each of the 8 combinations listed in Table 3-2, 5 balancing market simulations are run. The first of five wind actualizations is the historical forecast. The results shown in the next section analyze the average and the minimum and maximum resulting simulation.

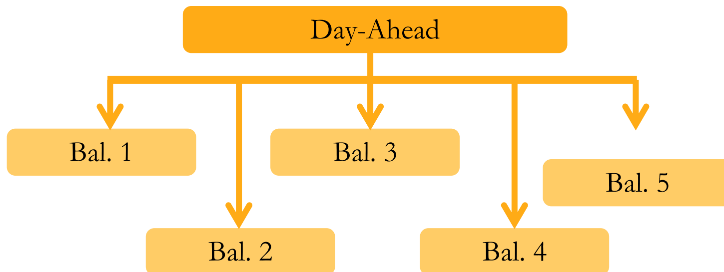

Figure 3-4 Simulations between day-ahead and real-time

3.4 SIMULATION DATA AND RESULTS

Each of the eight market design scenarios described in Table 3-2 are simulated for the day-ahead and balancing market with the goal of determining which improvement or set of improvements has the greatest benefit. The day-ahead market is simulated daily in one-hour increments for one year (365 days) using scaled historical wind and load data for each country using COMPETES. One day is optimized at a time, first for the day-ahead time-frame using the formulation in 3.3.2 and then for the balancing time-frame in which the realized wind value is different from the forecast. The schedule and commitment from the last hour of each day is used as a starting point for the following day in a rolling fashion. The balancing market takes the schedule from the day-ahead market and fixes the slow generators, allowing the fast-start generations to be committed. The balancing market is simulated using multiple real-time wind scenarios.

Between the day-ahead and real-time models, the wind forecast is updated but the load remains the same. This is due to several characteristics of wind forecast error compared to load error. First, load forecast error is usually much smaller than wind forecast error [9]. There are many factors that determine both types of error, such as wind farm size, location, and time horizon. Second, wind forecast errors are usually larger in range compared to load forecast errors, when expressed as a fraction of the forecasted amount [126]. Finally, the focus on this project is on wind uncertainty and integration, and this simulation allows me to focus on the impact of wind error without confounding it with load errors. There can be impacts of wind and load forecast error together, and indeed operating reserves should be acquired considering both, but data on both types of errors were not readily available for all countries.

The model is coded using AIMMS version 4.6 software using CPLEX 12.6.1 solver. The Europe-wide 24 hour model has approximately 145,000 constraints and 178,000 variables with 3672 integer variables for the day-ahead market. The Europe-wide balancing market has approximately 138,000 constraints and 130,000 variables with 3168 integer variables. Integer variables are only for units in the Netherlands, where there are many gas plants, which can be turned on during the balancing market due to their short minimum run times. Every day solves in 6-9 seconds, and including time to write output files, the yearlong simulation takes approximately 2 hours to run.

3.4.1 Characteristics of Generators, Wind, and Load

The characteristics of the generation in COMPETES can be found in the figures below, in addition to descriptions in [117], [118]. The generation mix used in the simulations is projected forward for the year 2030, which is the year considered in the simulations, and includes the current generation mix less planned retirements and any generation capacity already planned to open in future years. There is no additional capacity expansion modeled. Figure 3-5 shows the capacity for each country by fuel type. The gas generation includes gas turbines, combined cycle gas turbines, combined heat and power, and derived gas internal combustion plants [129]. Generators with a long minimum run time are considered slow-start generators and cannot be dispatched in the balancing market. In this study, all coal, lignite, nuclear, and biomass generators are slow generators, with the remaining being fast start generators able to turn on in the balancing market.

The wind capacity is a combination of onshore and offshore projected future capacity based on estimates for future penetration with data acquired by our collaborator, ECN's Wind Energy Unit. Capacity factors for wind are shown in Figure 3-6, for both onshore and offshore capacity. Histograms of forecast errors between day-ahead and balancing are shown in Figure 3-7 for the Netherlands and in Figure 3-8 for Germany. The errors shown in the figures are simple differences between the two data series, as was done in [126]. The root mean squared error for each of the five balancing wind simulations is shown in Figure 3-9 for each country as a fraction of the forecast output. An hourly average of the five wind series is shown in Figure 3-10 for Germany and the Netherlands. The amount of load in each country is shown in TWh in Figure 3-11, and load factors are shown in Figure 3-12.

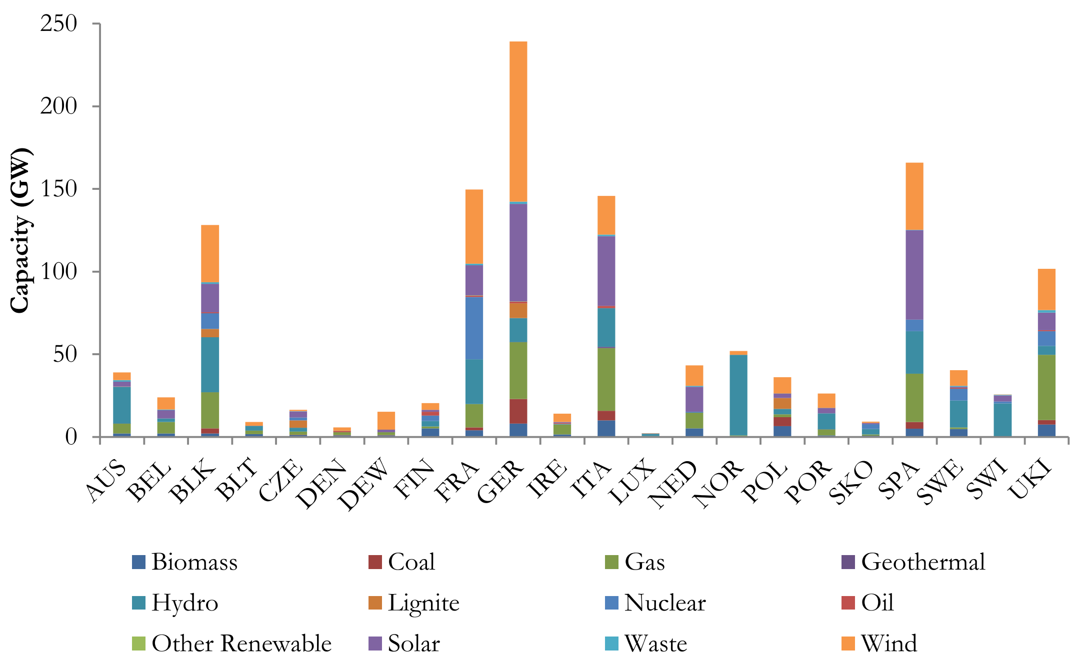

Figure 3-5 Generation capacity by fuel type in COMPETES for 2030 based on ENTSO-E Vision 4

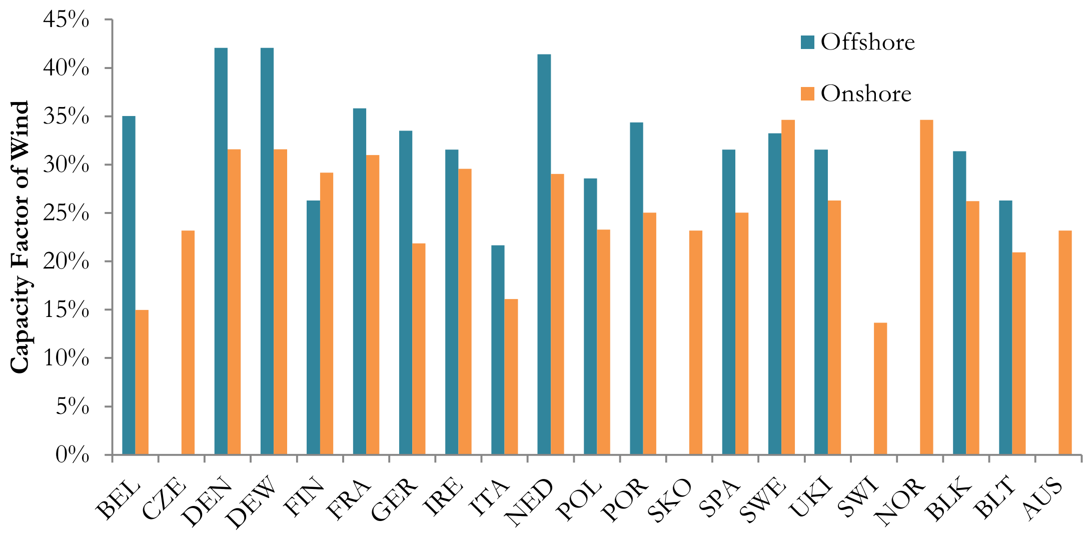

Figure 3-6 Capacity factor of wind for onshore and offshore

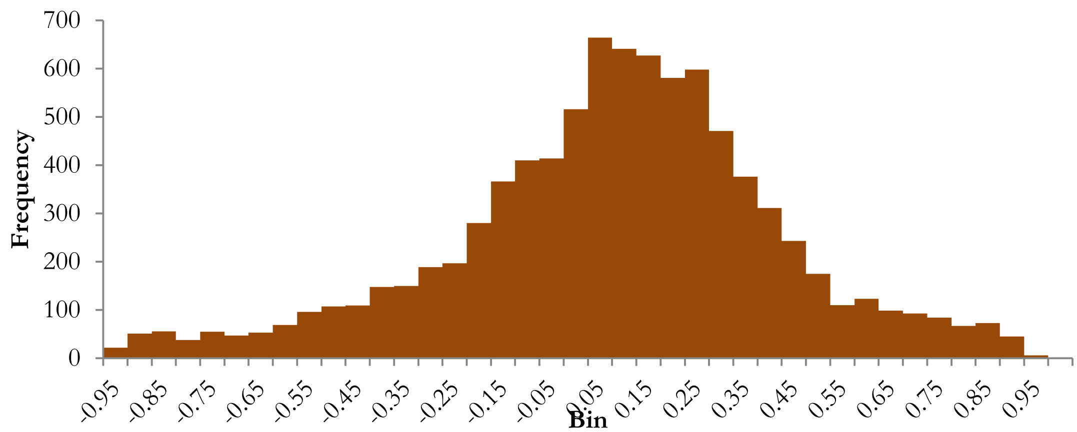

Figure 3-7 Histogram of onshore wind forecast error for the Netherlands as a fraction of the forecast output

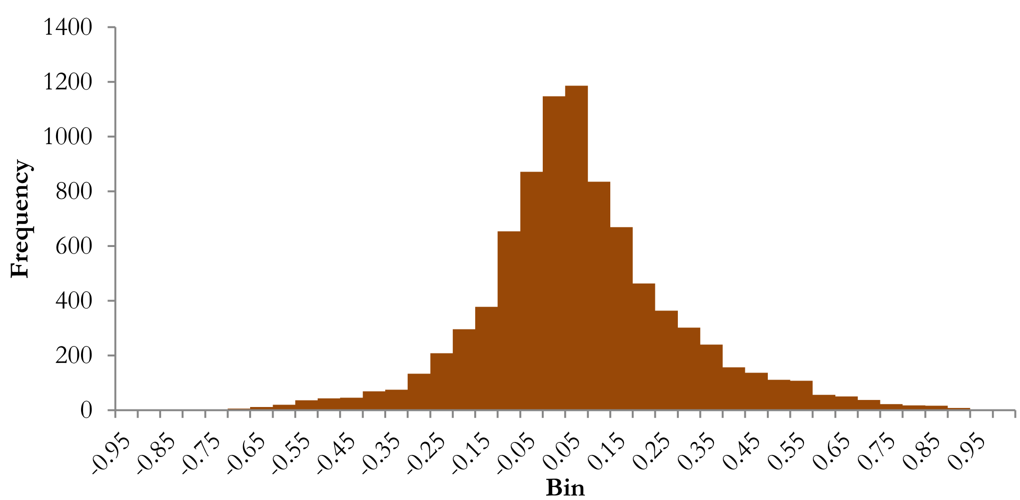

Figure 3-8 Histogram of onshore wind forecast error for Germany as a fraction of the forecast output

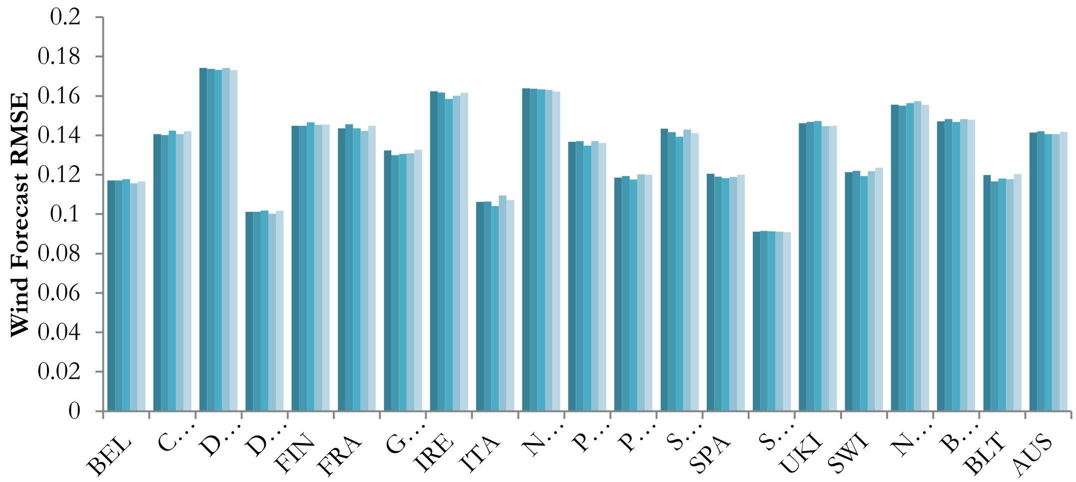

Figure 3-9 Root mean squared error for wind power between the day-ahead forecast and balancing actualization for each country, where each bar is a different balancing market annual scenario

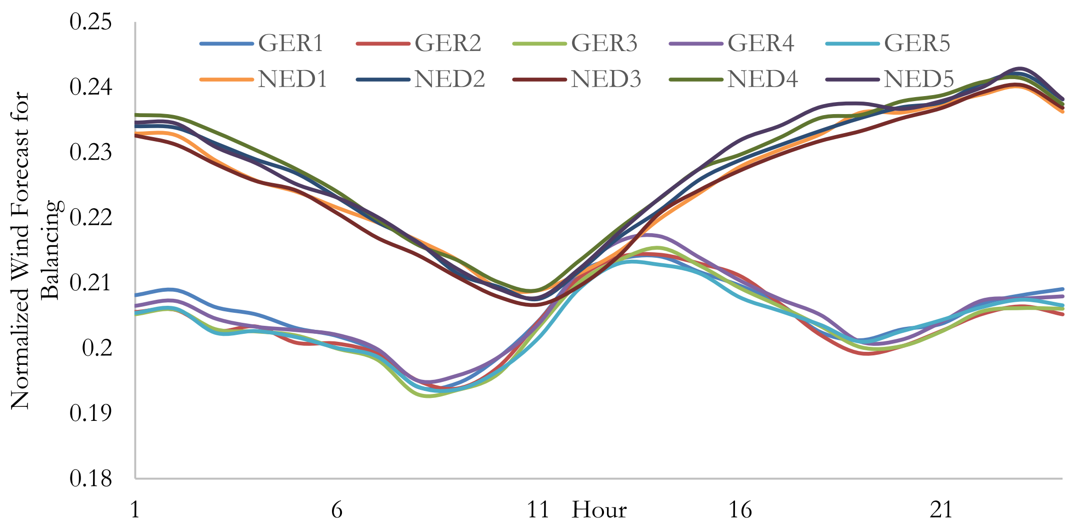

Figure 3-10 Five wind scenarios averaged by hour for one day in the Netherlands and Germany

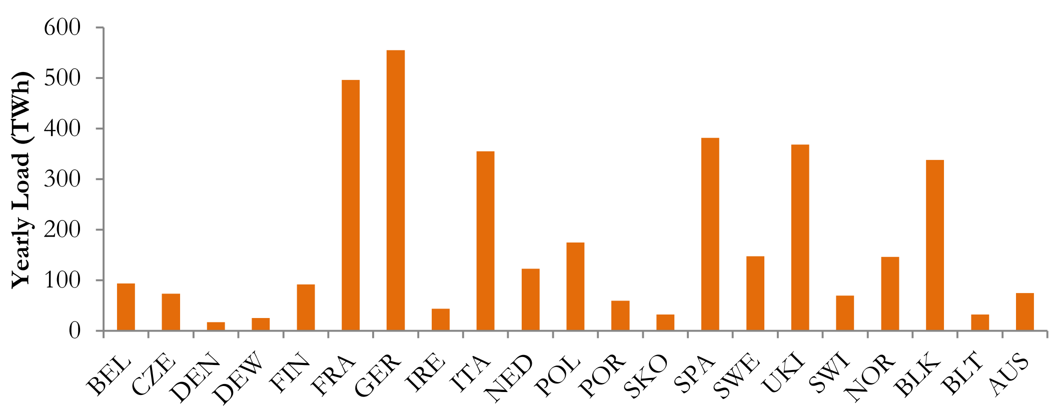

Figure 3-11 Load for each country for the study year, 2030

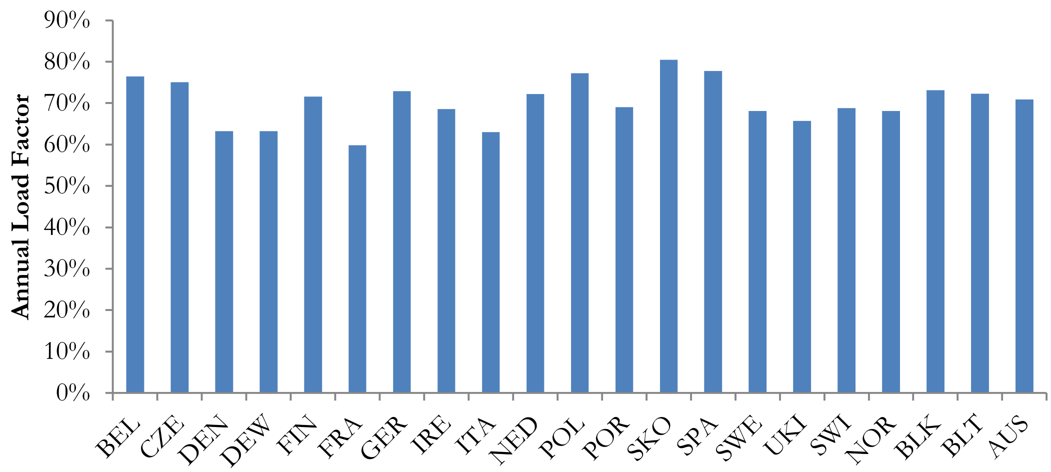

Figure 3-12 Annual load factor for each country

3.4.2 Operating Costs

The first assessment of market efficiency is examining total operating costs.[23] The operating costs from the eight simulations are shown in Table 3-3 and Table 3-4 as percent deviations from the hypothesized best-case simulation result; the same results are shown on the conceptual diagram in Figure 3-13. The value in the lower right corner of each table shows the 'ideal' result, i.e., the result with all three improvements simulated. Because the focus of the simulations is relative improvements between the improvements, the costs are shown as relative deviations from the ideal result. For example, the simulation using market-based, seasonal requirement with coordination in balancing only (MSCBal) is calculated using total operating costs as follows. The values shown are the output total operating costs for all of Europe for the yearlong simulation in millions of Euros.

$$\frac{MSCBal - ideal}{ideal} = \frac{\text{M€}90073.88 - \text{M€}89961.58}{\text{M€}89961.58} = 0.001248$$

All remaining values are calculated using this same relationship, with the 'ideal' case always as the comparison. The values in parentheses below the percentage are the minimum and maximum

[23] Since demand is inelastic, market surplus can be represented as a minimization of operating costs. If demand were elastic, the market surplus would be composed of both consumer surplus and producer surplus. In this model only the latter can be explicitly calculated.

deviations for total operating costs for each of the five annual simulations, also using the same calculation above. A non-parametric statistical comparison of results is shown later in the section.

Table 3-3 Total system cost (Europe-wide) without load shedding

(Left columns show allocation type and requirement type and top columns show the amount of coordination. Minimum and maximum are shown in parenthesis below each mean value.)

(Percentage)		No Coordination	Balancing Only	DA and Balancing
Contract	Seasonal	+110% (108, 111)	+30.0% (28.7, 30.8)	N/A
Market	Seasonal	+32.1% (30.9, 33.0)	+0.12% (-0.21, 0.51)	+0.11% (-0.23, 0.49)
	Daily	+40.9% (39.9, 41.7)	+0.05% (-0.03, 0.44)	0% (-0.33, 0.37)

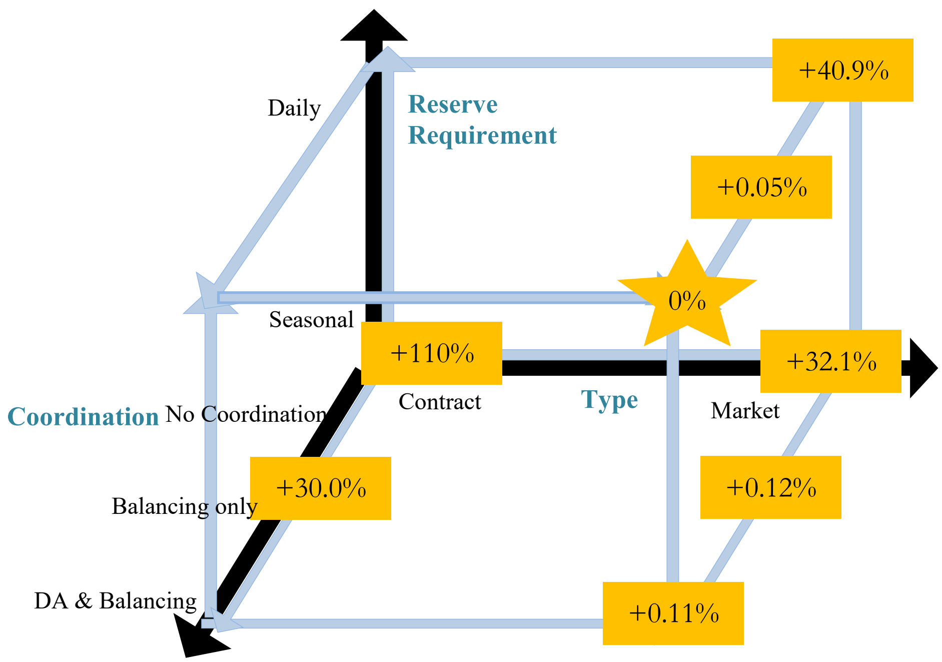

Figure 3-13 Conceptual diagram showing total system costs (Europe-wide) without load shedding

Table 3-4 The Netherlands system cost without load shedding (average percentage across five scenarios, with range across five scenarios of 365-day averages in parentheses)

		No Coordination	Balancing Only	DA and Balancing
Market	Seasonal	+29.1% (27.7, 31.4)	+0.67% (-0.48, 2.50)	-0.04% (-0.23, 0.49)
	Daily	+35.7% (34.3, 37.9)	+0.28% (-0.20, 0.91)	0% (-0.66, 0.95)

Table 3-5 Total system cost (Europe-wide) with load shedding (average percentage across five scenarios, with range across five scenarios of 365-day averages in parentheses)

(Percentage)		No Coordination	Balancing Only	DA and Balancing
Contract	Seasonal	+181% (179, 183)	+40.9% (39.78, 41.62)	N/A
Market	Seasonal	+55.4% (54.9, 56.1)	+0.14% (-0.15, 0.47)	+0.17% (-0.13, 0.51)
	Daily	+54.9% (53.9, 55.6)	+0.04% (-0.26, 0.39)	0% (-0.29, 0.32)

There are several distinct trends resulting from both Table 3-3 and Table 3-4. First, the ideal case is the lowest cost solution, i.e., all other percentages are positive. This confirms the hypothesis that the combination of all improvements produces the lowest cost solution. While this is true for the averages, the minimum and maximum costs for the scenarios overlap; meaning, the highest cost simulation from the balancing only coordination scenario for a market-based seasonal average is higher than the lowest cost simulation for the coordination in day-ahead and balancing scenario. This overlap hints that the operating cost results for certain scenarios might not be significantly different from each other.

Previous work of mine in collaboration with KU Leuven investigators [92] shows that a case with coordination in balancing alone has lower costs compared to a simulation with coordination in both day-ahead and balancing. The reason for the discrepancy is due to the availability of transmission capacity; the reserves allocated in one country in the day-ahead market could not be delivered in real-time because of lack of transmission capacity between the countries. The simulations in [92] have four countries with four transmission lines, whereas this set of simulations models all European countries with many interconnections. The results between the two coordination scenarios are much closer in this project, likely due to increased system size and availability of transmission capacity.

The overlap among ranges of scenarios leads to the second trend, which is evident from comparing the means of the no coordination cases with the two coordinated cases. The coordinated cases have much lower operating costs compared to the cases without coordination;

there is a minimum of 30% difference between the simulations. However, the difference between the two coordinated simulations is small, insignificant when considering the overlapping minimum and maximum scenarios. For example, the market simulation based on the seasonal requirement/coordination in balancing only assumptions has a minimum simulation that results in lower total costs (-0.21%) compared to the mean ideal case (but not lower than the minimum ideal case, -0.33%). This trend demonstrates that there is a great benefit in coordinating, either in activating reserves in the balancing market or both in the day-ahead and balancing markets.

The third trend can be seen in the difference between the seasonal and daily requirement results. There is a similarly small difference between the requirements for the coordinated cases: 0.12% compared to 0.05% and 0.11% compared to 0%. This difference is evident in the no coordination case when total system costs include load shedding. While the requirement was significantly different in these cases, it did not make a large difference in total system costs.

The five annual simulations performed for the balancing market do not provide enough yearly data for traditional parametric tests; therefore, a nonparametric test is used to analyze the yearly results. The Wilcoxon-Mann-Whitney (WMW) test is used to test if two samples are from the same population [130], [131]. The test ranks results from two populations, with the null hypothesis being the populations are the same and the alternate being that they are different [132]. Using the WMW test, the results can be analyzed using a z-score. The results for total European operating costs are shown in Table 3-6 as a comparison between the ideal case; the values are z-scores determined using the WMW test, where a score higher than 1.65 is outside the 95% confidence interval. For those scenarios with values greater than 1.65, the alternate hypothesis holds with 95% confidence, meaning, the samples come from different populations. For the simulations with a value less than 1.65, the null hypothesis holds with 95% confidence and the populations are indistinguishable.

The market-based simulations with some type of coordination are significantly likely to be from the same population. This indicates that doing coordination in balancing only or in both day-ahead and balancing could result in similar results. Simulations with no coordination are significantly different, and likely would not result in the low costs of the ideal simulation. The results for the total operating costs without load shedding in Table 3-7 show the same significance, where the seasonal simulations have a slightly higher z-score (still below 1.65). Comparing the two market-based simulations with no coordination against each other, the WMW test results in a score of 0.80, indicating they are likely from the same population. Comparing the market-based and contract-based no coordination simulations, the resulting score is 1.60; this indicates there is 94% chance they are from different populations.

Table 3-6 Z-scores using WMW test of Europe-wide system costs with load shedding; a score higher than 1.65 shows the annual costs are statistically significantly higher than the ideal solution (Daily market/DA and Balancing) with 95% confidence

		No Coordination	Balancing Only	DA and Balancing
Contract	Seasonal	2.61	2.61	N/A
Market	Seasonal	2.61	0.94	1.15
	Daily	2.61	0.52	Base Comparison

Table 3-7 Z-scores using WMW test of Europe-wide system costs without load shedding

		No Coordination	Balancing Only	DA and Balancing
Contract	Seasonal	2.61	2.61	N/A
Market	Seasonal	2.61	1.36	1.36
	Daily	2.61	0.52	Base Comparison

Although the overall lowest cost solution was the ideal case, each country's lowest cost solution was not in that scenario. Table 3-8 shows where each country had the lowest cost solution. Due to the similarities between the seasonal and daily simulations, these distinctions are likely not significant; meaning although the Netherlands is in the seasonal requirement, fully coordinated scenario box, it can be grouped with the daily requirement countries. The biggest distinction is the countries which are better off without coordination in both the day-ahead allocation and balancing activation phases. These countries generally see increased costs due to coordination because their generation is being ramped up for reserve or they are used for wheeling purposes. For the former case, a less expensive generator in Belgium might be used for balancing in surrounding countries. This would mean that surrounding countries reduce costs, but Belgium's costs increase since it is reserving more power in day-ahead. In a theoretical coordinated setting, the system will be at least the same or better off with coordination, but individual players in each country can be worse off. Generators might make more profit, but consumers might pay more for power. This can be seen in the statistical test for the Belgian system costs in Table 3-9. The seasonal balancing only simulation and the seasonal day-ahead and balancing coordination simulation are significantly different than the daily day-ahead and balancing simulation.

Table 3-8 Lowest cost solution by country

		No Coordination	Balancing Only	DA and Balancing
Market	Seasonal	BEL, FIN, IRE	-	NED, BLT
	Daily	DEN, POL, UKI	-	CZE, DEW, FRA, GER, ITA, POR, SKO, SPA, SWE, SWI, NOR, BLK

Table 3-9 Z-scores using WMW test for the Belgian system cost without load shedding; a score higher than 1.65 shows the results are statistically different with 95% confidence

(Percentage)		No Coordination	Balancing Only	DA and Balancing
Contract	Seasonal	2.61	2.61	N/A
Market	Seasonal	2.61	2.61	1.57
	Daily	1.77	0.94	Base Comparison

3.4.3 Trade

The biggest impact seen from comparing operating costs is the simulations comparing coordination of reserve between countries. Figure 3-14 shows the percent change between the ideal case and the seasonal/ market-based/ no coordination case, or the hypothetical 'worst' case of the market-based simulations (herein referred to as the worst case, which is at the origin of Figure 3-3). The total amount of energy from the balancing market is summed throughout the year and then compared against the other simulation. For instance, if the Netherlands exported 5 MW in hours two and four and imported 3 MW in hours one, three, and five, the total exports would be 10 MWh and imports would be 9 MWh. A positive value in the figure indicates that the ideal case has more trade, while a negative value indicates the worst market-based case has increased trade. For instance, Belgium both imports and exports more energy in the worst case compared to the ideal case, which is not unexpected when considering its lowest cost solution is also for this scenario (although the same is not the case for the other two countries that are better off in this scenario). Some countries, like Sweden and Denmark East, import significantly more when coordinating reserves, while Ireland exports more. Overall, the gross imports and exports in GWh increase by 0.16% in the ideal case compared to the worst.

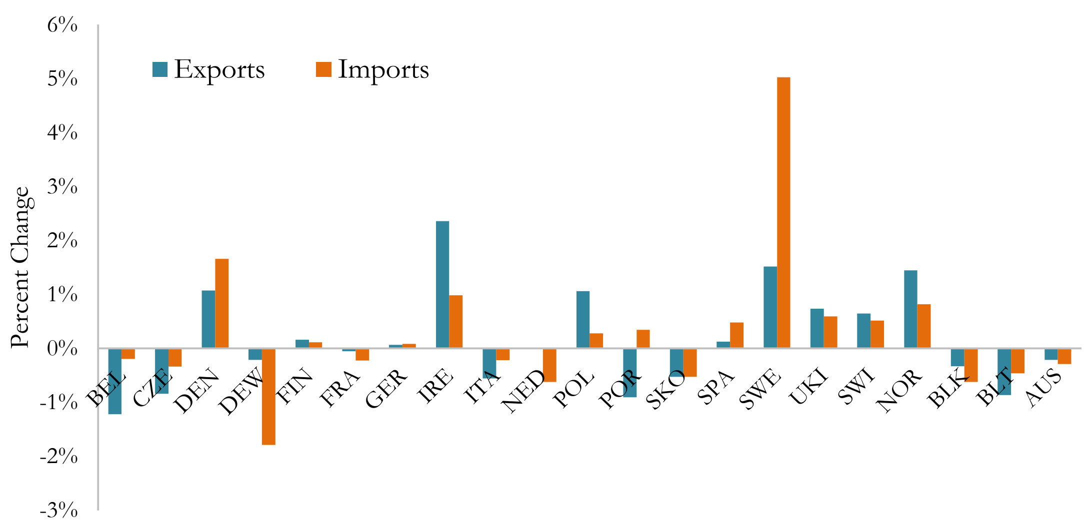

Figure 3-14 Percent change of total imports and exports between the ideal case and the seasonal requirement/ market-based/ no coordination case (worst market-based case)

3.4.4 Wind Energy

The wind and solar resources used in COMPETES are based on historical data for forecasted and realized wind, and modeled as free to the system operator. The historical data is normalized and scaled to expected wind penetration based on the European Network of

Transmission System Operators for Electricity (ENTSO-E) Vision 4 for renewable penetration [133]. The input wind penetration to the day-ahead model is shown in Figure 3-15, with the Netherlands at 30% penetration. Although the wind data is fed as a time series of input, the model allows for curtailment. Although the wind data is fed as a time series of input, the model allows for curtailment. As partial motivation for improving reserve procurement is lowering the cost of integrating wind power and improving its utilization, assessing wind curtailment is an important scenario evaluation tool. Curtailment percentages are Table 3-10 and Table 3-12.

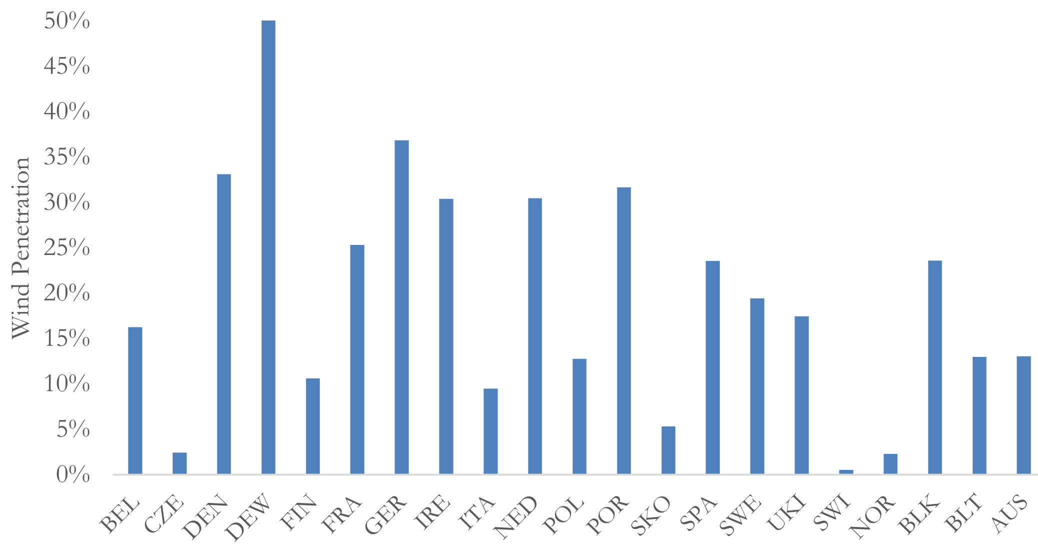

Figure 3-15 Wind penetration in GWh as a percentage of demand by country for a year

Table 3-10 Wind Curtailment in Europe

		No Coordination	Balancing Only	DA and Balancing
Market	Seasonal	+6.6% (1.7, 20)	-2.7% (-7.7, 10)	-0.6% (-5.3, 11.4)
	Daily	+6.2% (1.1, 19)	-3.5% (-8.2, 9.1)	0% (-4.8, 13)

Noticeably, the ideal case does not have the lowest curtailment of all scenarios. The daily requirement/ coordination in balancing only/ market-based scenario has the lowest overall curtailment, at -3.5% compared to the ideal case, followed by the seasonal case at -2.7%. In these scenarios, each country allocates its own reserves in the day-ahead market, meaning they might be more dispersed across the region compared to the ideal case, where reserve allocation is shared between countries. Then in the balancing market, reserve is spread throughout the countries, so more is available for balancing wind variability. The daily or seasonal fully coordinated cases might concentrate reserves in one region day-ahead, so that in balancing deliverability becomes difficult as lines are already delivering energy from day-ahead. As others have noted, reserving transmission line capacity can be critical to deliverability; Section 3.5 discusses how this might be done for COMPETES.

While the relative position of the simulations has changed, minimum and maximum curtailment overlaps between all scenarios. As seen in Table 3-11, the outcome of the WMW test show mixed results compared to the clear significance with operating costs. The two closest scenarios to the ideal case are likely from the same population, but the remaining simulations have over a 90% chance of being from a different population. The inconsistency in the curtailment results is partially due to low wind curtailment in the simulations; the total MWh of curtailment for the Netherlands is shown in Table 3-12, where there was no curtailment in four of the scenarios. Some countries saw no curtailment in many scenarios, making differentiation between simulations difficult.

Table 3-11 Z-scores using WMW test of Europe-wide wind curtailment

(Percentage)		No Coordination	Balancing Only	DA and Balancing
Contract	Seasonal	2.61	0.94	N/A
Market	Seasonal	1.78	1.57	0.52
	Daily	1.78	1.15	-

Table 3-12 The Netherlands wind curtailment (in MWh)

		No Coordination	Balancing Only	DA and Balancing
Market	Seasonal	3875 MWh (448, 8350)	0 MWh	0 MWh
	Daily	3885 MWh (448, 8350)	0 MWh	0 MWh

3.4.5 Operating Reserves

All thermal generation is available to provide reserve in the simulations. While pumped hydro storage is available to provide reserve, conventional hydro could not provide reserve. There are two types of reserve requirement examined in the simulations, a daily requirement and a seasonal requirement. Both requirements are based on the NREL 3+5 rule, meaning the average of 3% of demand and 5% of renewable forecast was reserved as backup power. The daily requirement used the average of the following day's load and renewable forecast, while the seasonal used the average of the entire season. The amount of reserve summed for each country in Europe for the daily requirement is shown by season in Figure 3-16 – Figure 3-19 as histograms. The seasonal average is described in the figure title. The daily requirements tend to have higher concentrations around the seasonal average, but show more variability, with secondary peaks in summer and fall.

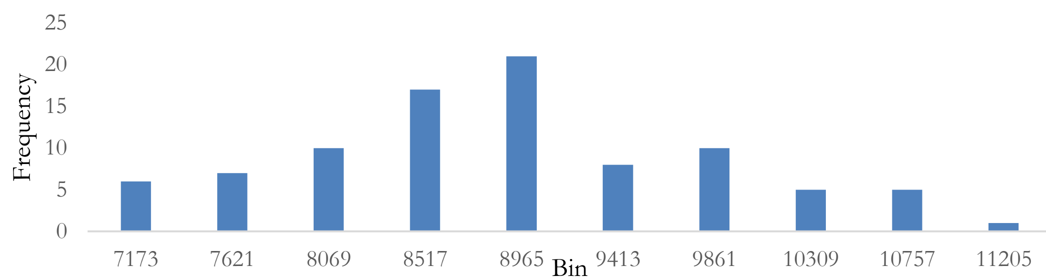

Figure 3-16 Dispersion of daily winter reserve requirements, winter seasonal average is 8675 MW

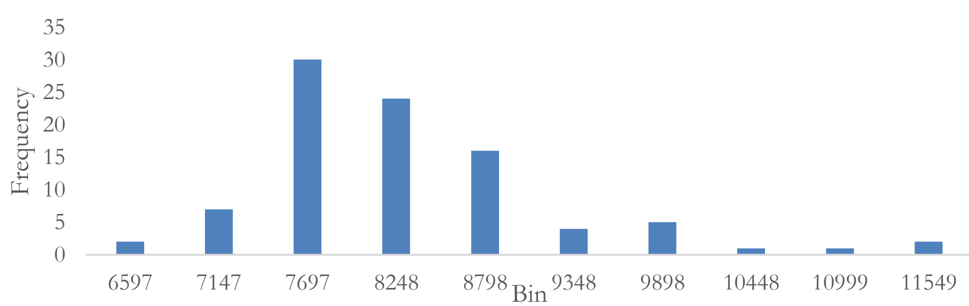

Figure 3-17 Dispersion of daily spring reserve requirements, spring seasonal average is 8032 MW

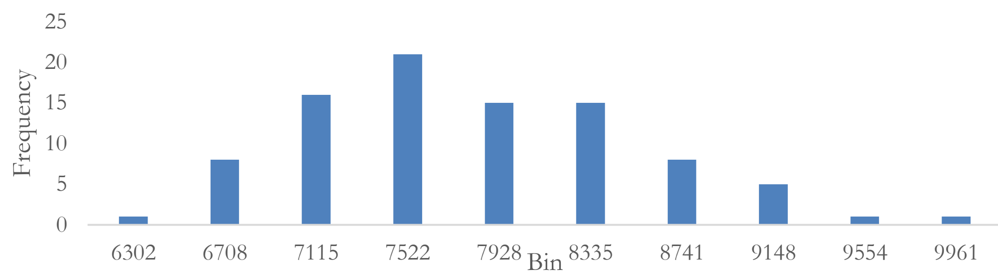

Figure 3-18 Dispersion of daily summer reserve requirements, summer seasonal average is 7625 MW

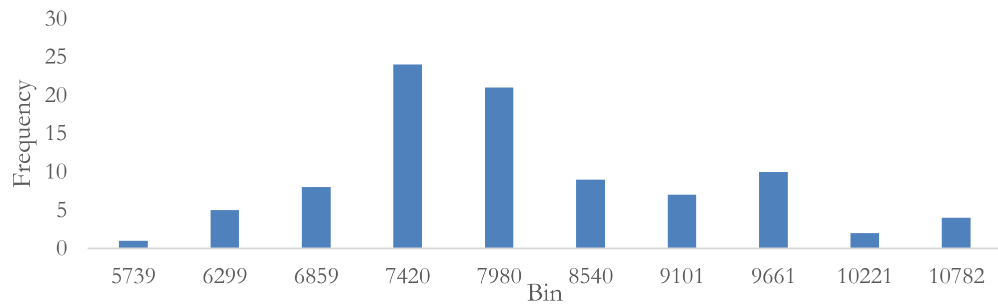

Figure 3-19 Dispersion of daily fall reserve requirements, fall seasonal average is 7870 MW

The hypothesis was that the daily average would be able to more closely approximate the needed amount of reserves, rather than consistently over- or under-estimating as a seasonal average would be. Yet, as Table 3-3, the choice of average time period in the reserve requirement had relatively little impact on operating costs. The cause of lessened impact is due to the amount of reserve available in the network. There reserve requirement constraint was rarely binding with a positive price, meaning there was usually more than enough reserve to support the requirement. Any extra headroom in a generator would be allocated towards the requirement, and there were enough generators throughout the network below their maximum capacity that it was not difficult to reach the requirement in most periods. This is partially due to the modeling of generation outside of the Netherlands; units are aggregations of several plants by fuel type and year installed. Therefore, the maximum capacity of the plants is overestimated, meaning this impact is likely underestimated. A comparison between the daily/market-based no coordination and complete coordination scenarios is shown in Table 3-13. The prices shown are averages of the nonzero periods, meaning if there were eight binding periods, the price shown is the average across the eight periods. There are generally fewer binding periods in the coordinated case compared to the case without coordination. For the Netherlands, Denmark West, and the United Kingdom, there are significantly more binding periods leading to increased costs. Reserve costs are largest in Sweden, being about €12/MWh on average; however, most countries have costs below €1/MWh.

Table 3-13 Comparison of average reserve price and binding periods between the daily requirement, market-based coordination scenarios

	Coordination in Day-Ahead and Balancing		No Coordination	
	Price (€/MWh)	Count of Binding Periods	Price (€/MWh)	Count of Binding Periods
BEL	0.003	2	0.023	28
CZE	0.002	1	0.004	2
DEN	8.017	4684	8.038	4701
DEW	0.003	2	9.218	5154
FIN	8.236	5064	8.25	5055
FRA	0.003	2	0.031	42
GER	0.003	2	0.004	2
IRE	0.388	215	0.481	278
ITA	0.006	8	0.006	8
NED	0.003	2	8.87	3046
POL	1.992	24	1.992	25
POR	0.074	61	0.079	72
SKO	0	0	0	0
SPA	0.074	61	0.079	72
SWE	12.153	7227	12.165	7215
UKI	0.003	2	0.476	275
SWI	0.002	1	0.002	1
NOR	0	0	0	0
BLK	0	0	0	0
BLT	0.046	53	0.047	54
AUS	0.003	2	0.002	1

The amount of reserve procured by generation type in the day-ahead market did not differ greatly between simulations and is shown in a single graph, Figure 3-20. Over half of the reserve was procured from storage, which is majority potential pumped hydro in the model. As mentioned in Section 3.3.2, the formulation used for storage reserves likely overestimates the quantity of reserves storage would provide. There is no clear consensus in the literature on the best way to model reserves provided from storage; future work would entail running the simulation with different reserve formulations to assess the overestimation. After storage, the next largest category is gas, which accounts for 37% of reserved allocation in the day-ahead market. Renewable resources were not chosen for reserve, since it was less expensive for renewables to be used for energy rather than reserve and existing thermal plants had enough headroom to account for the majority of reserve. The thermal plants used for reserve also included coal, lignite and a small amount of biofuels.

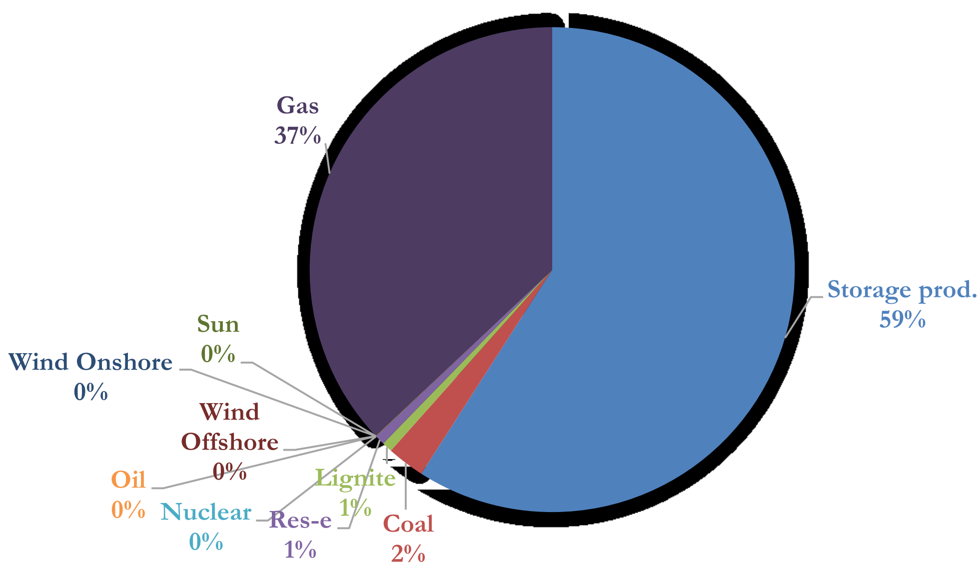

Figure 3-20 Reserve allocation by fuel source for year

3.5 DISCUSSION

With the challenges of uncertainty and variability, it is important for power operators to evaluate the operational flexibility of their system in order to cope with increased renewable generation. As the penetration of renewable energy increases, the system might require additional buffer capacity in order to accommodate the renewable sources. The extra capacity, operating reserve, can improve overall system reliability. Reserve procurement also impacts generators, who must commit reserve capacity well in advance of decisions about energy production. This reduces the efficiency of production and of reserve capacity provision. Allowing for simultaneous decisions about energy production and reserve allocation and deciding closer to real time reduces overall system costs. This chapter compares three categories of potential improvements to current practice of procuring, allocating and activating reserve in Dutch markets. Using a sophisticated pan-European energy market model, changes to reserve sizing, procurement methods, and coordination of reserve allocation are compared. If given the option to update the market using only one of the improvements, the results from this chapter can assist decision makers in choosing where to invest time and money. No other study has compared all three improvements simultaneously. As discussed in 3.2, studies have shown benefits of degrees of coordination, but none have also compared market- and contract-based allocation and different requirements.

3.5.1 Key Results

As hypothesized, combining all three improvements to operating reserve in a single model results in the lowest cost solution for Europe. Of the three improvements, coordination of reserve among countries in Northwest Europe[24] provides the single greatest improvement of the options. Both the market-based scenarios and the contract-based scenarios show greatest improvement when reserves were coordinated. Complementing the result of [92], coordination in the balancing market alone provides almost all of the benefit that coordination in the allocation (day-ahead) phase provides. By performing a small sensitivity analysis, the resulting simulation total costs are shown not to be significantly different from one another. The WMW test demonstrates that balancing market coordination lowers costs as much as coordination in both markets.

[24] The Netherlands, Belgium, Germany, France, and the United Kingdom

The results for coordination can be compared to previous studies of European markets. As mentioned previously, the comparison between different coordination schemes is also the central framework in [92]. Their results found that coordinating the allocation of reserves in the day-ahead market procured reserve that could not be delivered in real-time, resulting in higher costs. The balancing coordination alone was the lowest cost simulation. The results in this chapter show that the two cases are indistinguishable using the WMW test; one is not higher than the other, but both share the same benefits. This is likely due to the larger system and network size, since the network model used in this chapter connects the Northwest Europe to the Nordic region and southern European countries. Limitations of both models are discussed in the next subsection.

The difference in results between seasonal and daily reserve requirements is minor, with results not being significantly different in total costs between the five simulations. Although the daily requirement should have been a tighter bound, there was enough additional capacity in the system to provide sufficient reserve. This 'improvement' is likely not necessary to implement given the expected generation fleet. If the future fleet changes significantly, changes in the requirement might be necessary and further simulations can be performed.

Integration of wind is a major concern in the Netherlands and much of Europe looking forward. The results for wind curtailment point initially to the simulations with balancing coordination alone as having the least curtailed hours. However, the sensitivity analysis showed results varied widely across the scenarios, making it difficult to confidently suggest any one will dominate the others. Complementary research in the literature points to reservation of transmission line capacity as an important aspect of reserve delivery [92], [114]; future simulations can assess whether allocation of capacity could improvement wind integration.

The Dutch market regulators have been interested in improving aspects of reserve procurement. The results of the eight simulations suggest that the single improvement with the greatest impact is coordination of reserve in the balancing market. Rather than each country solely using its own resources in the balancing market, sharing resources across regions has the potential for significant cost savings. Coordination between countries is no small feat, as all European regulators would need to agree on trade rules and regulations. The advantage of this result is adjustments would only be required in the balancing market, rather than in both the day-ahead and balancing markets. As wind penetration increases across the Netherlands, higher levels of coordination can likely provide the Netherlands with both lower costs and greater levels of wind integration.

3.5.2 Limitations

Studying multiple time-frames using three improvements to current practice to simulate eight pairs of models can lead to limitations and suggestions for future work. The first improvement suggested, adjusting the reserve requirement from seasonal to daily, found little change in total system costs. Due to the type of reserve procured, the ramping headroom on the system was great enough to leave this improvement ineffectual. A future study should compare probabilistic requirements [103] or use of a response set [134].

Limitations might also be due to the type of reserved modeled. There are many products available in Europe and the U.S. and the need for each product can vary depending on generation mix. Future studies can compare contingency reserve, regulation reserve, and even ramping products[25] to determine which will most benefit the system. The modeling of storage might also be limited, since the framework used in this chapter will overestimate the resource's' ability to provide reserve. Other modeling constraints can be compared to identify the extent to which reserves have been overestimated.

The benefits seen in this chapter are also likely limited by the availability of transmission. Other studies have advised that deliverability of reserve might lower overall coordination benefits [112], [135]. Future simulations with this framework should analyze the impact of

[25] For instance, the California ISO has implemented a flexible ramping product [237], [238].

reserving transmission capacity for any cross-national trade. Finally, limitations of data availability for wind and load might skew results. This chapter also did not include load forecast errors, which might interact both with available of balancing reserve and wind forecast errors. Future studies should analyze the sensitivity of the results to correlations between wind and load and correlations between future wind plants in the region.

Generally, the results show large changes in operating costs between the different cases. These changes would be unlikely to materialize in the magnitude reported throughout the chapter (e.g., a 40% increase in operating costs). With the use of increased transmissions capacity, additional generation investment, correlated wind and load scenarios, and alternative reserve requirements, the magnitude is likely to decrease. Although the study has many limitations and suggestions for future enhancements, many insights are gained from comparing three improvements. This study can provide the building blocks for future analysis of reserve in the greater European market.

CHAPTER 4
PRICING ENVIRONMENTAL EXTERNALITIES IN REGIONS WITH ASYMMETRICAL POLICIES

4.1 INTRODUCTION

Managing the flow of electricity between adjacent electric grids is a challenging task for any system operator. Difficulties can arise when determining rules for trade, ensuring each side sees benefits, and coalescing contradicting policies. These difficulties are inflated when renewable energy and environmental concerns are involved. Wind and solar energy are prevalent in different parts of the country, and not necessarily co-located with load centers [136]. Trading that power across regional boundaries can create problems for system operators, who aim to maximize market surplus given local policies and rules on emissions reductions.

This chapter focuses on balancing area coordination considering the complicating factor of externalities, particularly, cases when neighboring balancing areas value an externality asymmetrically. An externality is a cost or consequence that is incurred outside of the market setting. There are both positive and negative definitions of externalities, which arise when the activities of one firm or person are dependent on another outside firm or person [137],[15]. Environmental externalities can be challenging to price [138]. If they are local, such as particulate matter pollution, the surrounding community is harmed. However, externalities like acid rain or carbon emissions from power plants do not stay local and are not necessarily created locally; they impact the region and even global communities [139]. If one balancing area aims to effectively reduce its carbon footprint, it must consider the effect of its decisions on net imports and exports and their carbon consequences; it is not enough to solely examine the power plants in the target region. A full accounting of a carbon footprint considers impacts of consumption in one region on emissions outside the region. There is rich literature on how carbon policies impact the electric system [140]–[144], which is discussed in detail in Section 4.2. The challenge becomes preventing leakage and contract shuffling [145]–[147], i.e., ensuring that a region's environmental rules do not take too much credit for reducing emissions because of unaccounted emissions increases outside the region. Even further, the challenge is to design policies that attempt to incent reduction of emissions in regions that are sources of imports.[26]

The Western U.S. is an ideal case study for examining the effect of power trading among regions that value emissions differently. California has implemented a cap-and-trade system for carbon emissions under AB 32 [26] while also completely consolidating its real-time market with several surrounding balancing areas through its Energy Imbalance Market [148]. Although several states have proposed legislation on carbon emissions, no other Western states have implemented a carbon reduction scheme. However, resources in those states still profit from selling power to California customers. Several proposals for coordinating trade have been proposed while attempting to maintain the integrity of California's AB 32 system (i.e., count the carbon emissions due to California imports while incenting carbon reductions outside California) [27], [149], [150]. This chapter evaluates each of the proposals for its ability to reduce carbon emissions, maximize market surplus, and price energy consistent with economic incentives. The

[26] While this chapter focuses on importing regions, exporting regions can also have similar concerns. A region may want to count emissions reductions that occur outside its region because of its exports. In the 1990s, Seattle City Light considered carbon reductions from exported power to California and found some planning scenarios had negative net carbon emissions due to exported power [239].

proposals are compared against a first-best (social cost minimizing) model and a do-nothing (no carbon regulation) model.

Broadly, this chapter delves into issues that can arise when some systems value a particular product but neighboring systems do not. Particularly, this chapter addresses the situation in which adjacent regions within a single market treat externalities in an asymmetric manner. The goals of economic efficiently and environmental regulation differ between regions, and improper pricing of the externality (the carbon reduction policy) might weaken environmental gains. The section answers the questions:

In an integrated market, can one region price an environmental externality that is not valued by all regions, and what inefficiencies might result?

What method for clearing the short-run energy market for multiple regions can account for carbon emissions that are due to demand in one region without causing cost-ineffective dispatch and distorted prices in another?

In particular, this chapter considers California's efforts to lower carbon emissions under AB 32 [26] while still encouraging power trade with states elsewhere in the West that do not limit emissions. This chapter contributes to current literature through modeling and analytical comparison of different greenhouse gas emission schemes. The market model simulations evaluate proposals for incorporating carbon emissions, finding no one method dominates the others in both reduction of costs and emissions. Section 4.2 reviews literature on carbon policies, such as cap-and-trade, incorporation into markets, and the broader institutional context for greenhouse gas modeling. Model descriptions can be found in Section 4.3, along with detailed mathematical formulations. The do-nothing and social cost of carbon models are in Section 4.3.1, followed by the single-pass method in Section 4.3.2, the two-pass method in Section 4.3.3, and the tax-at-the-border method in Section 4.3.4. The results are shown in Section 4.4, including explicit definitions of carbon leakage in Section 4.4.2.e, followed by a discussion in Section 4.5. In the discussion, I consider the broad set of criteria that are relevant to choice of a policy to limit leakage of carbon emissions from a regulated region to a neighboring unregulated region that has no carbon regulation, and describe how the results of this chapter shed light on some but not all of those criteria.

4.2 BACKGROUND AND LITERATURE REVIEW

Coordination of energy and environmental markets in regions that exchange electricity is difficult. There can be differences in policies and procedures, in addition to distinct political motivation. Inevitably, any two regions might have different policies when it comes to operation of the electric grid. Since there is no federal policy for greenhouse gas emissions in the U.S., individual states have implemented their own emissions reduction policies. Although this section focuses on the carbon policy in California, it is applicable to many regions that are considering an emission reduction scheme. The sub-section describes the cap-and-trade system, with background literature on the strengths and weaknesses of different implementation options. The second sub-section defines of carbon leakage and options for a second-best carbon modeling. Finally, the third sub-section includes further details on the history of cap-and-trade in California and current practices.

4.2.1 Options for Valuing Emissions under Cap-and-Trade Regulation

Coordination when regions employ different policies is a complex issue, one that might not have a clear solution. One issue that has become especially prominent due to climate change is policies around greenhouse gas emissions. Different states in the U.S. have enforced policies around the amount of greenhouse gas emissions that can come from the power sector. In the Northeast, the Regional Greenhouse Gas Initiative (RGGI) mandates resources within its member states buy allowances for greenhouse gas emissions, with a total cap on emissions updated over time [151]. In California, the California Air Resources Board (CARB) has put in place a similar allowance system, requiring all resources selling to the California market purchase allowances under Assembly Bill 32 [26].

These systems are called cap-and-trade and have been used in the U.S. and Europe to curb emissions. Cap-and-trade limits the total emissions in a region by giving generators allowance credits that can then be traded in a given time period [152]. By allowing trade, individual generators can decide how much they can emit depending on the price of the allowance. While the focus of this chapter is not on the efficacy of cap-and-trade or other renewable incentive programs, literature on the efficacy is extensive for the U.S. [143], [153], [154], Europe [141], [142], [155], between-countries [156], and compared to other renewable incentive programs [12],[157]. Some major issues that are addressed are the ways in which allowances are allocated to participants, its impact when competing with other renewable incentive programs (such as Renewable Portfolio Standards (RPS)), and leakage (discussed later in this section). From a policy standpoint, the focus of this chapter is on the impact of an existing cap-and-trade scheme within one region on multi-region electricity markets, rather than the scheme itself.

Assuming a region is implementing a cap-and-trade system, there are several decisions that must be made to account for emissions; broadly, these have been divided into load-based, source-based, and first seller approaches [11], [146], [152], [158]. A load-based system forces the demand-side to account for emissions from the power it consumes. If the load exclusively uses bilateral contracts to buy power, the load-based system would be easy to track. However, in the current electricity market framework, identifying the exact source of emissions is extremely difficult [159] or not feasible [158]. Additionally, since inter-state sales can occur, further complications arise from tracking emissions from these sales. Through simulation, [160] shows load-based systems could raise costs and inhibit competitiveness. Further showing the shortcomings of the mechanism, [161] shows load-based systems have increased transactions costs, being at best equivalent to source-based systems.

The source-based system focuses on accounting for emissions from the supply-side. As was described above in reference to RGGI, allowances are allocated to individual sources of emissions; these can then be bought or sold depending on need. This method is the most popular and can work well in regions where all participants are subject to the cap-and-trade scheme. The resource with the allowance will add any extra costs into their energy bid. In this way, they can recover the cost and the supply stack will be reordered to account for resources with new higher costs. However, this method becomes difficult when multiple regions are involved; for instance, California buys power from participants outside the state who are not subject to AB32. To deal with this issue, [11] designate "pure" and "modified" methods, where the pure source-based approach would exclude imports and the modified version includes power sold to California (or the target region).

Lastly, cap-and-trade can be implemented using a hybrid method called the first-seller approach. Like the source-based approach, the first-seller scheme assigns emissions to the supply side for generators within California. For imports, the entity importing the electricity is assigned the emissions responsibility. Since the first-seller approach can account for emissions due to in-state generation and imports, it is preferred by many including explicit endorsements from the authors of [146] and [160]. Formal analysis of the method, called first-deliverer in the paper, is simulated in [162], where the authors find that it is not likely to be more effective than the source-based approach, and only slightly reduces emissions when certain rules are in place.

These three major methods address the source of emissions accounting under a cap-and-trade regime. If utilized in a market context, each method will have distinct implications for prices and schedules output from the market. To analyze the impacts of both emission trading and market interactions of the different accounting methods, two papers distinguish themselves by considering an endogenous allowance price [163], [164]. The first focuses on the method and equity of allowance allocation [163]. By modeling two allocation methods, the authors assess how each will impact leakage and profitability. They show that an allocation method providing high emission rate plants with more allowances to prevent cost shocks ("fuel-based" updating) might instead increase costs and emissions in the long run relative to allocating allowances based on production output ("output-based" updating). The second paper focuses on the strategic behavior of generators under a carbon trading scheme [164]. They find generators with high emissions rates are incentivized to take a long position in forward markets, driving up prices, and allocation of allowances to generators with low emissions rates would decrease prices. The results are complementary; both find that allocating more allowances to high emission rate plants would not necessarily reduce emissions or prices. Both papers' authors also acknowledge their modeling departs from realistic markets, as they use a modified version of source-based allowances. Even with this assumption, each draws important implications for markets.

Both simulation and regression can be used to analyze the impact of cap-and-trade on electricity prices. Simulation has been used more extensively used for analysis of European cap-and-trade system, called the European Emissions Trading Scheme (ETS) [165]–[167], whereas cap-and-trade in U.S. affects fewer markets [168]. Engineering economic simulations anticipate, consistent with economic theory, that at least some of the cost of carbon allowances would be passed through to consumers, with the exact amount depending on price elasticities of supply and demand, and the exact generation mix [11]. Regression analysis for the European ETS examined whether futures electricity prices would respond faster to increases or decreases of CO_2 allowances and fuel prices, finding that there was no evidence of asymmetric responses [169].[27] Although earlier work found no impact on electricity prices in California [170], the same authors in [171] found the price of carbon in California significantly impacted several surrounding regions. They advocate for expanding the cap-and-trade region, as the prices are already being impacted by the program. This analysis underscores the difficulty of the California

[27] The authors specifically analyzed the asymmetric response of prices to increases and decreases in input prices (such as fuel), called "rockets and feathers."

cap-and-trade efforts. Assigning allocations and creating energy market rules for recovery of the allocation price is exceedingly difficult when regions have asymmetric policies.

Much of the literature on cap-and-trade has focused on implementation of the cap-and-trade auction and allocation methodology. Similar to [169]–[171], the research of this chapter begins with the assumption the auction has occurred, and a price for emissions has been determined. Instead of analyzing if carbon prices will influence electricity prices through regression, this work simulates different methods for integrating carbon prices into generator bids. It focuses on the issues that arise when neighboring regions participate in the electricity market, but are under no legal obligation to reduce carbon output locally. This complication, as is present in California, creates opportunities for divergent incentives and price manipulation. Through simulating multiple bidding methodologies, this chapter compares costs, prices, total emissions, and leakage to determine if one method dominates the rest. The next section will describe existing research and definitions of leakage, followed by a discussion of the institutional context for this chapter.

4.2.2 Emissions Leakage and the Second-Best Response

Leakage can occur when neighboring grids have asymmetrical policies[28] [25], [162], and is defined as increased emissions outside of the target area due to imports [172]. Meaning, even if emissions decreased in the target region (e.g., California), they might have increased elsewhere in the network (e.g., Southwest or Northeast). Leakage can be calculated simply by comparing the emissions in nonregulated region under a carbon reduction scheme to baseline emissions from a market with no carbon regulation. Leakage can be expressed as a fraction:

(Increase in nonregulated region emissions)/(Decrease in regulated region emissions).

[28] Leakage can occur on both the supply and demand side [162]. Demand-side leakage occurs when consumption of a good decreases in the target region and increases elsewhere. Long term supply-side leakage occurs when plants move out of the target region in response to increased regulation.

Literature on emissions trading schemes commonly uses this definition of emissions leakage [28]–[31], [172]. It is most relevant to cap-and-trade systems that only apply to the regulated region and do not attempt to attribute emissions to imports or regulate them (e.g., RGGI, Clean Air Act SO_2 Trading Program, and European ETS). Leakage using this definition can occur if energy imports from the non-regulated area to the regulated area increase and if the regulations provide no incentive or ineffective incentives to limit measures in the non-regulated area.

There are variants on this definition; one proposed in this chapter suggest that if a region's policy accounts for estimated emissions associated with imports from another region, then leakage can be defined as a discrepancy between the emissions accounted for in imports, and actual changes in emissions outside the region due to imports. Since carbon emissions are not localized, policies that aim to reduce carbon emissions should consider regional impacts rather than local areas alone.

A related issue that arises is contract shuffling [173], in which power importers change the designated source of power from dirty to clean sources without actually changing the operations of facilities [146]. As a hypothetical example, California places value on reducing carbon emissions by use of allowances, while Idaho, Nevada and Utah do not value limiting carbon emissions. Say that a California utility presently has a power purchase contract with a Utah coal plant. If the California policy penalizes imports that are associated with high emitting sources, then the utility could switch its contract to existing hydropower plants in Idaho and Nevada. The utilities in Idaho and Nevada that formerly bought hydropower could then obtain replacement power from the coal plant in Utah. Utah can sell any kind of power to those utilities, including power that has high carbon emissions. Although California is buying nominally clean power from Nevada, Utah is still producing power from the high carbon emitting plant, meaning the emissions are still taking place, just not directly being sold to California. Dispatch and physical power flows are unchanged by the policy, but the emissions accounting system inaccurately reports that emissions are reduced. This is an example of contract shuffling.

The first-best response to capturing carbon leakage and contract shuffling would be a tax on all externalities in the region. In the case of California, it would mean taxing all power being dispatched by the ISO within California as well as other states. This would be the best method for implementation since all generators would both bid their allowances and the prices would reflect the cost. However, the surrounding states have not implemented a cap-and-trade program, meaning they do not value greenhouse gases in this way. The first best solution would therefore unduly tax their customers, and it is unlikely any state would support such a policy. This chapter analyzes what method might be second-best [174], in that it can achieve carbon reductions cost-effectively subject to the constraint that California cannot impose a tax or shadow price on all carbon emissions by power plants in other states. Any method that is second-best cannot capture the full extent of carbon emissions, but comes close given the policy and regulatory backdrop of the region [144].

While the modeling in this section can be applied to any externality that is subjected to differential regulation in neighboring jurisdictions, the language and descriptions in the section focus on California's challenge of attaining an emissions goal while integrating with neighbors. From an economic standpoint, there is an environmental externality that must be internalized into prices in one region and not affect prices in another.

4.2.3 Institutional Context

California and its neighbors have begun full coordination (integration) due to the great benefits both sides are expected to achieve [148]. This market, called the Energy Imbalance Market (EIM) [175], is a voluntary real-time market in the Western U.S. that co-optimizes all participating resources as if they were one balancing area. Utilities and companies in surrounding states have opted to bid their resources into the California market and abide by CAISO market rules, but the market does not extend to include their residents in California state policies. These different regions have diverging priorities, where California alone has a policy goal of reducing greenhouse gas (GHG) emissions. Even if large economic benefits are achieved from complete coordination among real-time markets, there is a concern that emissions targets will not be met [176].

Concern over greenhouse gas emissions arose as the CAISO was attempting to determine a new governance structure for a West-Wide ISO [177]. The original proposal had a specific principle related to GHG emissions; however, it was removed in the final proposal and noted it was not directly related to governance [178]. It was later developed as an issue paper [179]. This is not a new concern for California, but the models used to address the issue have changed over time. One of the methods examined in this chapter (the single-pass method) and current practice for California can result in carbon leakage, a concern for environmental groups and the California Air Resources Board [31], [180].

If plants bid their allowances into the real-time market, the lowest cost resources will be dispatched for each load, likely dispatching high cost resources last. However, in a single optimization, here called the single-pass method, leakage and contract shuffling are still issues. In this chapter I compare the results of the single-pass method with other real-time market clearing procedures that have been proposed to lessen the amount of leakage and/or contract shuffling. One such procedure was proposed by the CAISO [148], [149], essentially establishing a counter-factual that allows California to attribute emissions to imports it buys from other states, and more accurately characterize the emissions that can be attributed by its own load. Due to criticism, such as [27], the proposal was recently amended, which is one optimization model (rather than two passes)[29]. Although the model has changed and likely will continue to change, a comparison can still provide insight into the relative differences of these methods. The results of a two-pass system can be compared against other approaches to avoid leakage, particularly a tax-

[29] The proposal was amended in February 2018, and this dissertation was completed in April 2018.

at-the-border method, which places a hurdle rate on trades, limiting emissions but not capturing all leakage.

The question of bidding allowances into energy markets is relevant to several ISOs at present. California went through an extensive stakeholder process to develop the two-pass methodology (described in Section 4.3.3) [148]. Due to the rapid developments in the politics of regional coordination and the emergence of new approaches, especially in the California context, academic literature is just beginning to address the topic.[30] The author in [27] points to several disadvantages of the two-pass proposal and briefly lays out an alternative. As stated in the paper, the core shortcomings of the two-pass proposal are its deviation from efficient dispatch, creating perverse bidding incentives, increasing the demand for carbon permits. As an alternative, the paper suggests imposing an exogenous fee and using a separate settlement structure that would maintain efficient dispatch and charge each party according to regional policy.

Outside of [27], several reports have been published that discuss the incentives and implications carbon pricing on markets. In response to New York's carbon goals, NYISO commissioned a report to analyze carbon pricing for the state [23]. The extensive report addresses alternative carbon pricing options, market design issues such as leakage and allocating carbon payments back to customers, and the overall benefits to markets. They identify two main ways to create carbon prices for New York: add a carbon price for each MW generated or create a secondary (and tighter) cap-and-trade system. While each has its benefits, neither can fully apply to the case in California, since California dispatches generators in states not under the cap-and-trade system. Although the New York market is separate than its neighbors, it is part of RGGI, where there is an embedded carbon cap.

[30] The effort to reform integration of greenhouse emissions into the California market began in September 2016 with CAISO's announcement of a new initiative, "Regional Integration California Greenhouse Gas Compliance" [179]. After a straw proposal was published in November, a draft final proposal was issued in June 2017, followed within the month by a revised version [149]. The second revision was published in February 2018 [240].

Unlike New York ISO, the PJM system operator is composed of states with differing carbon policies. PJM is evaluating options to manage carbon in the region, and have published a report most similar in scope to Section 4.2.1 [25]. They analyze both the two-pass method proposed by CAISO and a single pass method (described in Sections 4.3.3 and 4.3.2 respectively); while neither method is strictly preferred, they simulate an example where the two-pass method reduces emission further than the single-pass.

The simulations in this chapter similarly analyze both the two-pass and single-pass options. However, the work in this chapter extends the comparison to a tax-at-the-border alternative and rigorously evaluates the resulting prices, dispatch, payments, and costs. No report or paper to date has analyzed the variety of carbon pricing methods for wholesale markets given asymmetric carbon polices, as is the case in California. The simulations offer a framework for assessment of existing and proposed mechanisms, and provide further guidance on how carbon costs incorporated into market algorithms can impact cots and emissions in a region.

Each of these models will be described in the next subsections, followed by examples. The models refer to California in the formulation, because it is the target case study and to ease understanding. It is also consistent with the cited formulations for the CAISO's proposed two-pass model. This methodology, however, can be generalized for any two or more sets of regions. However, further complications might arise if any other surrounding state implements a carbon tax or cap-and-trade system, such as the one Washington put on the ballot in 2017. This issue will continue to arise in regions with adjacent balancing areas, such as PJM [25], New York [23], and the European Union.

4.3 GHG TRADING MODELS

In order to examine the impacts of different carbon pricing schemes, five approaches are modeled and analyzed. Two approaches represent the extreme cases: no greenhouse gas pricing and a socialized price for all. These two models will bookend the three cases that incorporate carbon pricing and attempt to account for carbon embodied in imports using different types of models: a single-pass, a two-pass, and a tax-at-the-border method. Mathematical models for each of the five methods are explained in the subsections below. Two methods are modified from California ISO's modeling efforts to integrate greenhouse gas trading into their market [148]: single-pass and two-pass methods. The CAISO has changed the model formulation over the last few years and very recently decided to adjust the method to a single pass, different from the one shown in this section. The stakeholder process is ongoing and negotiations with the California Air Resources Board are still to come. While the method discussed will not be implemented, it still provides a counterpoint for the remaining models.

The five models will be compared based on region specific and total emissions, costs, surpluses, and prices. The focal outputs, emissions and costs, are compared in a two-dimensional plot to see how the five simulations compare. A sensitivity analysis will then be presented to determine the sensitivity of the results to different the chosen carbon prices.

4.3.1 Upper and Lower Limits: No Greenhouse Gas Model & Social Cost of Carbon

At the upper and lower limits of the model types are the no greenhouse gas (No-GHG) case and the model which includes a social cost of carbon. The No-GHG model should have high emissions, as cost is the only factor in considering dispatch; it should be the upper limit for emissions and the lower limit for cost. The bookend to a No-GHG model would be a 'full' GHG model, or one where each and every generator considers the social cost of carbon. It would be the lower bound for emissions, and near the upper bound for cost. Although this is not the only method to determine a lower emissions solution, out of the methods proposed below, it is the first-best solution to lowering emissions because all generators face the same price of carbon, and carbon has the same impact no matter where and when it is emitted. The method is only a theoretical possibility, unless all regions and generators are mandated (or agree) to dispatch considering the same cost of carbon. However, it is a useful benchmark against which the emissions reductions and costs of other solutions can be compared.

There are several constraints that are the same for all five models, and are common to most economic dispatch models. The common constraints, together with an objective that involves only fuel and other variable operations & maintenance costs and excludes carbon penalties, create the 'No-GHG' case, where plants are dispatched solely based on their characteristics and marginal costs. The objective in (4-1) is simply minimizing the cost of operations, or marginal costs multiplied by quantity dispatched. Lower and upper bounds on generation and line flow are in (4-2) and (4-3). Flow on a line is calculated using the Power Transfer Distribution Function method (4-4), and (4-5) is the node balance constraint. For simplicity, and to be consistent with previous chapters, D_i^{max} represents the total inelastic (fixed) demand for the period, rather than a variable for demand dispatched.

No Greenhouse Gas Model Formulation

$$\text{Min} \sum_{\forall t} \sum_{\forall g} C_g p_{g,t} \tag{4-1}$$

Subject to

$$0 \le p_{g,t} \le P_g^{\max} \qquad \forall g, t \tag{4-2}$$

$$F_k^{\min} \le f_{k,t} \le F_k^{\max} \qquad \forall k, t \tag{4-3}$$

$$f_{k,t} = \sum\nolimits_i PTDF_{i,k} \left(\sum\nolimits_{g \in G_i} p_{g,t} - D_{i,t}^{max} \right) \qquad \forall k, t \tag{4-4}$$

$$\sum_{g \in G} p_{g,t} - \sum_{i \in N} D_{i,t}^{\max} = 0 \tag{4-5}$$

Mathematically, a model with no carbon price and one with a socialized price are similar. They differ only in the bid in cost of the generating facilities. The Social Cost of Carbon method would change the objective to include the GHG cost bid along with the marginal cost bid, as in

(4-6). Otherwise, the model remains the same. This formulation is used as a first-best comparison, showing possible outcome if all regions valued carbon the same way.

Social Cost of Carbon Formulation

$$\text{Min} \sum_{\forall t} \sum_{\forall g} (C_g + C_g^{GHG}) p_{g,t} \tag{4-6}$$

Subject to

(4-2) – (4-5)

4.3.2 Single-Pass Method

In order to capture the power flow due to power plants emitting greenhouse gases, constraints and variables must be added to record or tag plants that emit greenhouse gases. The single-pass method allows the system operator to limit the imports of power into California to those plants who bid a greenhouse gas cost and quantity into the market. The greenhouse gas costs are accounted for separately from other costs for generators outside California and included in a single energy bid for generators within California. Within California, they are one and the same, as the costs will be paid by California residents. Generators outside of California also have the option of only serving their local load by submitting a quantity bid of zero to the California market (i.e., their power can't be used to support exports to the CAISO); CAISO will then dispatch them in merit order without any sale to California load. With a non-zero capacity bid, the resource can provide up to that amount of power to California customers. By using greenhouse gas cost bids in addition to energy bids, prices in California will reflect combined energy and greenhouse gas costs. Outside of California, prices are calculated without a greenhouse gas adder, so residents of other states do not pay for California state regulations.

The single-pass method suffers from the contract shuffling problem described in the literature review and introduction to Section 4.2.1, in which greenhouse gas emitting plants in Utah might displace renewable energy in Nevada thereby maintaining rather than reducing emissions. This issue is often referred to as contract shuffling [145]. In a single pass method, the power being sold to California will have lower emissions, but does not impact sales among the surrounding regions selling.

The model formulation from CAISO's June 23, 2017 model [149] is used as the basis for the single-pass method, and the two-pass method in the next section. In the single pass method, the objective has two sets of terms. Like the economic dispatch model above, the marginal costs of production are minimized. In addition, we must assign a cost to GHG emissions. As discussed in Section 4.2, generators are required to purchase allowances to offset the emission of CO_2 from an emission auction, and this value will determine the GHG price bid that is offered into the real-time auction. In addition to the price bid, a new quantity for production from a plant whose GHG emissions are attributed to California consumption must be introduced. The variable captures the net flow of power to California from greenhouse gas emitting plants, $p_{g,t}^{GHG}$. The new bid is multiplied by the new quantity and minimized in the objective in (4-8), below.

The remainder of the additional constraints defines the bounds on the GHG quantity variable. The lower bound of the sum is defined in (4-9) by the total power generated outside California less demand, i.e., anything imported into California. The upper bound for each generator is defined in (4-10)-(4-12) as the minimum of the power dispatch of the generator, its capacity, or the GHG quantity bid into the market. In actual operations, these models would be the final auction before power delivery. Before this model runs, there would be a fifteen minute, an hour-ahead, and day-ahead market auction. The results of those auctions would be part of the input to this model. Since this comparison is only considering a single market framework, additional inter-auction constraints are omitted.

For both the single-pass and the two-pass models, prices (LMPs) outside of California include an adder for greenhouse gas emissions. If the dual variable on the node balance constraint (4-3) is λ_t, the transmission constraint (4-5) is $\mu_{k,t}$ (+ for upper bound and – for lower bound), and the greenhouse gas lower bound (4-9) is η_t, then prices at each node are defined as in (4-7). The system energy price () can reflect additional costs of greenhouse gas emissions. By adding the greenhouse gas lower bound dual, η_t, the prices outside of California will only reflect the marginal energy price and the additional costs of greenhouse gas allowances. This ensures that load only pays the marginal cost of energy and not the additional greenhouse gas bid.

$$LMP_{i,t} = \lambda_t + \sum\nolimits_k PTDF_{i,k}\mu_{k,t}^{+} - \sum\nolimits_k PTDF_{i,k}\mu_{k,t}^{-} + \eta_t \qquad \forall g, t \qquad (4\text{-}7)$$

Single-Pass Formulation

$$\text{Min} \sum_{\forall t} \sum_{\forall g} (C_g p_{g,t}) + \sum_{\forall g \notin CA} (C_t^{GHG} p_{g,t}^{GHG}) \qquad (4\text{-}8)$$

$$\sum_{g \notin CA} p_{g,t} - \sum\nolimits_{i \notin CA} D_{i,t}^{\max} \le \sum_{g \notin CA} p_{g,t}^{GHG} \qquad \forall t \qquad (4\text{-}9)$$

$$p_{g,t}^{GHG} \le p_{g,t} \qquad \forall g, t \qquad (4\text{-}10)$$

$$p_{g,t}^{GHG} \le G_{g,t}^{GHG} \qquad \forall g, t \qquad (4\text{-}11)$$

$$p_{g,t}^{GHG} \le P_g^{max} \qquad \forall g, t \qquad (4\text{-}12)$$

and (4-2) – (4-5)

4.3.3 Two-Pass Counterfactual

The two-pass system was designed in an attempt to address the leakage problem that can occur with emissions trading in the single-pass method. As described earlier in Section 4.2.1, a single pass method might apparently reduce emissions directly sold to the interested party (in this case California), but maintain or even increase emissions elsewhere in the network. In an attempt to limit the substitution of external polluting generation for within-California generation (i.e., leakage), so that total emissions are reduced, a two-pass method was proposed by the CAISO. In order to limit the substitution, the first pass must establish the baseline emissions that would occur without California. This is done by limiting the first pass to trade outside of California. The schedules from the first pass are used as input for the second pass, but prices from the first pass are not used for settlements. The second pass then uses the baseline to limit emissions; the upper bound for generation capacity becomes the difference between total capacity and the optimal baseline from the first pass. The two passes of the two-pass system are described below.

First Pass

The first pass limits imports by California to be non-positive in total, so that California must supply all of its demand, and the remaining nodes can trade amongst each other or buy California exports. This is enforced through (4-14), where the sum of the flow on lines going into California is non-positive, allowing exports from California to the remaining nodes. Although nodes outside California will not directly pay for greenhouse gas emissions, the greenhouse gas dispatch variable is included in the first pass to set a baseline for the second pass. This enables the model to distinguish between emissions caused by California and those that would be emitted otherwise.

As in the single-pass, the GHG dispatch variable is included in the objective only for the generation outside of California, seen in (4-13). The lower and upper bound for $p_{g,t}^{GHG}$ is defined as it was in the single-pass method and the trade is limited GHG capacity submitted to the market, shown in (4-15) and (4-17)-(4-19). The only time $p_{g,t}^{GHG}$ will be positive is if the previous markets pass on generation from one region. In the model this value is shown as T_{15}, which represents the fifteen-minute market. In all simulations, this value is zero.

Two-Pass Counterfactual: First-Pass Formulation

$$\min \sum_{\forall t} \sum_{\forall g} \left(C_g p_{g,t}\right) + \sum_{\forall g \notin CA} \left(C_t^{GHG} p_{g,t}^{GHG}\right) \tag{4-13}$$

$$\sum_{\forall k \in CA} a_k f_{k,t} = 0 \qquad \forall t \tag{4-14}$$

$$\sum_{g \notin CA} p_{g,t} - \sum\nolimits_{i \notin CA} D_{i,t}^{\max} \leq \sum_{g \notin CA} p_{g,t}^{GHG} \qquad \forall t \tag{4-15}$$

$$\sum_{g \notin CA} p_{g,t}^{GHG} \leq -T_{15} \qquad \forall t \tag{4-16}$$

$$p_{g,t}^{GHG} \leq p_{g,t} \qquad \forall g \notin CA, t \qquad (4\text{-}17)$$

$$p_{g,t}^{GHG} \leq G_{g,t}^{GHG} \qquad \forall g \notin CA, t \qquad (4\text{-}18)$$

$$p_{g,t}^{GHG} \leq P_{g}^{max} \qquad \forall g \notin CA, t \qquad (4\text{-}19)$$

and (4-2) – (4-5)

Second Pass

Using the first pass GHG dispatch as a baseline, the second pass can redispatch generation throughout the region considering cost and limited by the baseline emissions. There are two modeling differences in the second pass. The limit imports to California is eliminated, and the limits on $p_{g,t}^{GHG}$ are now tighter. The difference between the maximum capacity, and energy and GHG dispatch from the first pass now creates an upper bound for the GHG dispatch variable. This limits the dispatch so that any change from the first pass can be attributed to demand from California. In this way, the 'leaked' power will now be influence the price and final energy dispatch.

A brief example can illustrate how the two-pass method works. A generator outside of California has a maximum capacity of 100 MW and submits a greenhouse gas quantity bid of 90 MW due to previous local contracts of 10 MW. The first pass dispatches the generator to 60 MW. The GHG dispatch variable is the lesser of the capacity (100 MW), the GHG quantity bid (90 MW), and the energy dispatch (60 MW); it happens that the optimal solution is 55 MW in the first pass. In the second pass, the GHG variable will be the limited again by the dispatch, GHG quantity bid, and the last term of (4-21), which is 100 MW – (60 MW – 55 MW) = 95 MW. Since the GHG quantity bid is lower than the last term, the variable would be limited to 90 MW.

Two-Pass Counterfactual: Second-Pass Formulation

$$\min \sum_{\forall t} \sum_{\forall g} \left(C_g p_{g,t}\right) + \sum_{\forall g \notin CA} \left(C_t^{GHG} p_{g,t}^{GHG}\right) \qquad (4\text{-}20)$$

$$p_{g,t}^{GHG} \leq p_{g,t} \qquad \forall g \notin CA, t \qquad (4\text{-}21)$$

$$p_{g,t}^{GHG} \leq G_{g,t}^{GHG} \qquad \forall g \notin CA, t \qquad (4\text{-}22)$$

$$p_{g,t}^{GHG} \leq P_g^{max} - \left(p_{g,t}^{*(1)} - p_{g,t}^{GHG*(1)}\right) \qquad \forall g \notin CA, t \qquad (4\text{-}23)$$

and (4-2) – (4-5), (4-15), (4-16)

4.3.4 Hurdle Rate or Tax-at-the-Border

The last method takes a simplistic approach: limit the amount of GHG emissions coming into California by putting a hurdle rate or tax at the border on all imports. Similar to the hurdle rate used in Section 5.5.3, this rate puts economic friction on any power flowing into California. The power will not be imported unless its value is above the tax [158]. Unlike the former hurdle rate, this value is asymmetric. Since only California values emissions, the import tax is much higher than the export, or the export tax is zero or negative. Unlike the previous methods, the tax-at-the-border method overlooks the individual emission rates of power plants. Any power coming across the border must 'hurdle' the tax, which can include plants for a variety of emission rates.

The formulation adds a term in the objective, (4-25), to tax the flow into California. Two new variables are introduced for these flows, s_t^{iCA} and s_t^{eCA} , respectively. The variables are defined in (4-26), and only apply to the intertie lines into and out of California. The input values chosen for the GHG import tax (T_t^{GHG}) can have a significant impact on the outcome of this model. The export fee (T_t^{export}) can be negative, maximizing the power leaving California; if the power is primarily from plants with higher emissions, total emissions might not decline overall. This simulation set the export fee to zero and varied the level of the import tax. The results show how different import taxes impact the results, and future simulations can examine how varying the export fee might impact outcomes. The prices that result from the hurdle rate formulation include the dual variable of the import/export (φ_t) constraint in (4-26). The calculation for prices at each node is in (4-24).

$$LMP_{i,t} = \lambda_t + \sum_k PTDF_{i,k}\mu_{k,t}^{+} - \sum_k PTDF_{i,k}\mu_{k,t}^{-} + \varphi_t \qquad \forall g,t \qquad (4\text{-}24)$$

Tax (Hurdle Rate) Formulation

$$\min \sum_{\forall t} \sum_{\forall g} \left(C_g p_{g,t}\right) + T_t^{GHG} s_t^{iCA} + T_t^{export} s_t^{eCA} \qquad (4\text{-}25)$$

$$s_t^{iCA} - s_t^{eCA} = \sum_{\forall k \in K^{\mathrm{IT}}} f_{k,t} \qquad \forall t \qquad (4\text{-}26)$$

and (4-2) – (4-5)

4.4 WESTERN U.S. EXAMPLE

4.4.1 Network Description

The three methods for GHG trade and carbon accounting for imports are compared in an example using the network and generation modified from [11]. The three-bus ten-generator example is a rough approximation for the Western U.S., where California includes GHG costs in their prices while the remaining two nodes, the Northwest and Southwest, exclude GHG costs from prices. The generator and load data can be found in Table 4-1 and Table 4-2, and the network configuration in Figure 4-1.

A single hour is modeled. The generators from [11] represent aggregated units, broadly mimicking the generation in each region. The Northwest has a great deal of hydro, which is represented by Gen 6 with zero GHG emissions and zero marginal cost. The remaining generation is not specific to a technology, but has representative characteristics; the Southwest has some lower cost resources with higher emissions compared to the lower emissions and higher costs of California. Compared to [11] the marginal cost, capacity, emissions rate, and network characteristics have not been modified. Since the GHG bid and cost are specific to this problem, they were calculated based on the provided emissions rate. The calculations multiplied the emission rate by the prevailing price for CO_2 allowances. The November 2017 allowance auction produced a rate of \$15.06/metric ton CO_2 [181]. To reflect the increasing trend of allowances, this project used a rate of \$17/metric ton CO_2, which is in line with other cap-and-trade schemes like RGGI [23] and in the same range as those used in simulations of California [170]. The GHG costs that result are in Table 4-1. This value can greatly influence the outcome of the different methods; a sensitivity analysis is done with results in Section 4.4.4. Additionally, the flow on the line between the Southwest and California is modified to increase available capacity to allow for further trade among all three regions.

Table 4-1 Generator characteristics

	Location	Marginal Cost (\$/MWh)	Capacity (MW)	CO_2 Rate (kg/MWh)	GHG bid (MW)	GHG cost (\$/MWh)
Gen 1	CA	28.14	250	580	-	9.86*
Gen 2	CA	26.46	200	545	-	9.27*
Gen 3	CA	26.6	450	600	-	10.2*
Gen 4	NW	15.52	150	500	150	8.5
Gen 5	NW	16.2	200	500	200	8.5
Gen 6	NW	0	200	0	180	0
Gen 7	SW	17.6	400	1216	400	20.67
Gen 8	SW	16.64	400	1249	400	21.23
Gen 9	SW	19.4	450	1171	450	19.91
Gen 10	SW	18.6	200	924	200	15.71

*These values are not used as GHG cost bids in simulations, rather are added to the marginal costs and bid as energy costs. The resulting California marginal cost bids total 38, 35.72, 36.80 \$/MWh respectively. Only the No-GHG case uses the marginal cost values from the table.

Table 4-2 Network characteristics

Line/Node			CA	NW	SW
	Max Capacity	Load	890 MW	303 MW	684 MW
Line 1	255 MW	PTDF Values	+0.3333	–0.3333	0
Line 2	120 MW		+0.3333	+0.6667	0
Line 3	60 MW		–0.6667	–0.3333	0

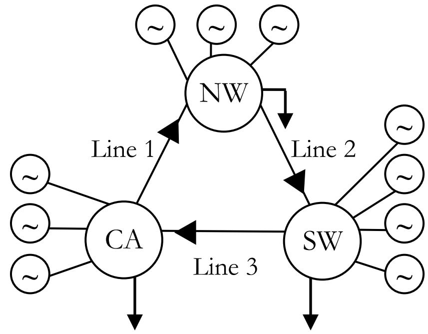

Figure 4-1 Three-bus network diagram

4.4.2 Simulation Results

Each of the five methods was modeled in AIMMS and solved using CPLEX. The GHG models have approximately 53 constraints and 23 primal variables, and the hurdle rate model has 23 constraints and 15 variables. The results are shown Sections 4.4.2.a-4.4.2.e, which describe market outcomes, including: dispatch, operating costs, emissions, prices and profits, and carbon leakage. In the results section, the term 'GHG models' refers to the simulations that include GHG costs in some form, including the single-pass, two-pass, tax, and social cost of carbon models (excluding the No-GHG case). The first pass of the two-pass system is shown in the results for reference, as the proposed two-pass method does not include a pricing run for the first pass.

The tax-at-the-border method also shows two results: below and above the effective break point. Properly selecting a tax or hurdle rate is one difficulty with this method, and examining many hurdle rates enables evaluation of both successful and unsuccessful rates. In this simple example, there is a break point above which the tax limits all trade to California and acts like the first pass of the two-pass method, thus making it ineffective. In this case, "ineffective" does not mean carbon emissions are reduced. It means the regional market is ineffective in terms of promoting trade (it discourages trade). Below this point, the tax takes effect and allows some trade between regions. For this example, the break is \$20.40/MWh; anything at or below

$20.40/MWh would allow trade at that price while anything above it would limit trade to nothing. One additional point lies with a tax of $0/MWh, where the results would be equivalent to the No-GHG case. Different problems will have different break points and determining that point remains a difficulty with the tax method. The break point for this problem is found by looping through incremental tax rates to find changing solutions. A summary of each model type can be found in Table 4-3.

Table 4-3 Summary of model types

Name		Description
No-GHG		No greenhouse gas costs in the objective for any generator (in California or the rest of the West (ROW), no additional constraints
GHG only in CA		Greenhouse gas costs in the objective only for California generator but not for ROW, no additional constraints
Single-Pass		Greenhouse gas costs part of CA generator's energy bid ("GHG in CA MC"), separate variables and constraints to account for and penalize greenhouse gas costs for ROW generators deemed to export to CA
Two-Pass	First Pass	No imports or exports to CA, sets a baseline for greenhouse gas accounting, GHG in CA MC
	Second Pass	Allows imports and exports, identifies and penalizes the greenhouse gas imports using baseline from first pass and same constraints as the single-pass, GHG in CA MC
Tax	Effective	Net flows into CA must pay a tax (or buy allowances) based on an assumed ROW-wide marginal emissions rate and tax rate/allowance price, equivalent to 20 $/MWh; GHG in CA MC
	Ineffective	Same as above except the tax (> 20 $/MWh for this case)
Social Cost of Carbon		Greenhouse gas costs included in all CA and ROW generators' energy bids, no additional constraints

4.4.2.a Dispatch

The total dispatch level is shown in Table 4-4. Both the No-GHG and the 'GHG only in CA' models result in the same dispatch, and are shown in the same column. The difference between these models is the carbon price on generators in California; in the latter model, generator offers in California are composed of the marginal and GHG costs added together. Given the input data in this chapter, the imports between the Northwest and Southwest region and California are particularly limited. The transmission capacity between the Southwest and California is small, preventing additional imports between the regions. The marginal (fuel) costs are also high in California, much higher than the rest of the West. When the GHG costs are added to the marginal costs, merit order does not change, meaning the supply stack stays the same.

Table 4-4 Power dispatch level in MW by plant, and net exports (generation minus load) by region

	Region	No-GHG/ GHG only in CA	Single Pass	Two-Pass First Pass	Two-Pass Second Pass	Tax Effective (<20)	Tax Ineffective (>20)	Social Cost of Carbon
Gen 1	CA	26.7	30	240	105.6	90	240	77
Gen 2	CA	200	200	200	200	200	200	200
Gen 3	CA	450	450	450	450	450	450	450
Gen 4	NW	150	150	150	150	150	150	150
Gen 5	NW	200	200	132.9	200	200	132.9	200
Gen 6	NW	200	200	200	200	200	200	200
Gen 7	SW	250.3	247	104.1	171.4	187	104.1	0
Gen 8	SW	400	400	400	400	400	400	400
Gen 9	SW	0	0	0	0	0	0	0
Gen 10	SW	0	0	0	0	0	0	200
CA net export		-213.3	-210	0	-134.4	-150	0	-163
NW net export		247	247	180	247	247	180	247
SW net export		-33.7	-37	-180	-112.6	-97	-180	-84

With No-GHG limits and no effective tax, California imports the greatest amount of power, followed by the single pass method and the effective tax. If California is required to supply its own power, in the first pass of the two-pass method or the ineffective hurdle rate, the high cost Gen 1 is used almost at capacity. In the remaining cases, Gen 7 increases output to supply California and Gen 1 is operated a lower level. Including the cost of emissions, Gen 7 cost slightly more compared to Gen 1 ($38.27 compared to $38). The single pass model backs down the units in 'clean' order rather than merit order, meaning it backs down the highest marginal and carbon cost unit rather than the highest marginal cost unit. The two-pass and effective tax models also back down Gen 7, but significantly more than the single pass method. Seeing only the marginal and carbon cost of Gen 7, the social cost of carbon model does not turn it on, and instead turns on Gen 10, which has a lower marginal and carbon cost.

The last three rows of Table 4-4 show the net exports, or generation less load, for each region. The imports into California are highest for the No-GHG case and the social cost of carbon case. The first pass of the two-pass model and the ineffective tax do not import or export power to California.

4.4.2.b Operating Costs

Next, we examine the total variable operating costs of each node. These are costs of fuel and non-fuel variable operations & maintenance and, as indicated below, either include or omit the expense of AB32 emissions allowances. The total costs are separated into three tables, Table 4-5 through Table 4-7: the first shows costs including GHG costs for California based on AB32 allowances and the taxed imports (for the tax model), the second includes all additional GHG costs due to AB32 allowances for all regions, and the third excludes GHG costs altogether. Note that Table 4-6 shows the true social cost, including all non GHG operating costs plus the social cost of carbon emissions. It can be obtained by multiplying each generator's emissions rate by the social cost, assumed to be $17/ton.

Note that the actual operating costs of any system can include or exclude GHG costs depending if the allowances procured by generators are given for free or auctioned. If allowances are free, then the costs in Table 4-7 would reflect operating costs.[31] If they are auctioned, and all generators outside of California are able to recover the costs or are allocated allowances for free, Table 4-5 shows the total costs per region. The tax columns in Table 4-5 show the tax rate in the

---

[31] This assumes that the generators who emit GHG and therefore must hold allowances are given them for free; if however allowances are allocated freely but to generators who do not need them and therefore sell them to generators, then the distribution of costs will differ, although the total in Table 4-7 would be correct.

costs for the Northwest and Southwest, adding a tax of $20.40/MWh to the imports from each region into California (imports totaling 150 MWh). The ineffective tax does not import into California, which is why the costs are higher in the Northwest and Southwest for the effective tax compared to the ineffective tax. If all regions require generators to purchase allowances, then Table 4-6 would reflect the generator's cost of operating under GHG pricing.

Unsurprisingly, the lowest cost solution in Table 4-5 and Table 4-7 (both of which do not include any GHG costs in that particular solution) is No-GHG, the case in which no plant pays for AB32 allowances. Its cost is shown as the same in both tables, because there are no carbon costs charged to any plant in that regulatory model. The second lowest cost solution is the single pass method, which trades the most power among the GHG methods, thereby operating the least cost generators. Next is the two-pass method followed by the effective and ineffective tax models. The tax payment is excluded from the totals in Table 4-7, since it would depend on the exact amount. As noted earlier, the ineffective tax has the same solution as the first pass of the two-pass system, allowing no trade and operating at the highest cost.

California's operating cost follow the same trend as total system costs, with the No-GHG solution as the lowest cost, followed by the single pass. The Northwest operating costs only differ when no trade is allowed to California; the remaining solutions are the same. Because their plants are operating at capacity, their dispatch is only impacted when all trade is blocked. Finally, costs in the Southwest increase under the GHG schemes, and become lower when trade is blocked. Because Southwest plants operate at higher output, total costs grow when allowed to trade with California. However, their revenues will also increase when they sell more. Since the marginal unit sets the price in the Southwest, profits happen to be the same for all simulations; further results are shown in Section 4.4.2.d.

The operating costs including the cost of carbon in all regions (as charged by each respective regulatory scheme) are shown in Table 4-6; for consistency with the remaining models, carbon costs are shown in Table 4-6 for the No-GHG case. As might be expected, the costs of the No-

GHG method including carbon costs are higher than the GHG models. The social cost of carbon model has, by construction, the lowest cost, since the carbon costs are considered during dispatch. The second-pass and the effective tax have similar solutions, and are second-best only to the social cost of carbon solution. Both models do not turn on Gen 10, which has a lower combined marginal and carbon cost, but are able to use more power from Gen 1 ($38/MWh combined) than Gen 7 ($38.27/MWh combined). The highest cost solution is the ineffective tax and the single pass, since both require the use of high cost resources in California.

Table 4-5 Total operating costs ($/hour) (excluding non-CAISO GHG costs)

	No-GHG	GHG only in CA	Single Pass	Two-Pass		Tax		Social Cost of Carbon
				First Pass	Second Pass	Effective (<20)	Ineffective (>20)	
CA	18,013	24,720	24,844	32,824	27,718	27,124	32,824	26,630
NW	5,568	5,568	5,568	4,481	5,568	8,264	4,481	5,568
SW	11,061	11,061	11,003	8,488	9,672	10,310	8,488	10,376
Total	34,642	41,348	41,415	45,793	42,958	45,699	45,793	42,574

Table 4-6 Total operating costs ($/hr, including GHG costs for all regions assuming social cost of $17/ton for CO_2)

	No-GHG/ GHG only in CA	Single Pass	Two-Pass		Tax		Social Cost of Carbon
			First Pass	Second Pass	Effective (<20)	Ineffective (>20)	
CA	24,720	24,844	32,824	27,718	27,124	32,824	26,630
NW	8,543	8,543	6,886	8,543	8,543	6,886	8,543
SW	24,728	24,602	19,133	21,708	22,306	19,133	22,011
Total	57,990	57,989	58,843	57,969	57,973	58,843	57,184

Table 4-7 Total operating costs ($/hour, excluding GHG costs for all regions)

	No-GHG	Single Pass	Two-Pass		Tax		Social Cost of Carbon
			First Pass	Second Pass	Effective (<20)	Ineffective (>20)	
CA	18,013	18,106	24,016	20,234	19,795	24,016	19,429
NW	5,568	5,568	4,481	5,568	5,568	4,481	5,568
SW	11,061	11,003	8,488	9,672	9,947	8,488	10,376
Total	34,642	34,677	36,985	35,474	35,310	36,985	35,373

4.4.2.c Emissions

The model with the highest emissions solution is the case without any GHG constraints, seen in Table 4-8. The single pass, effective tax and two pass methods follow respectively, each with lower total emissions. As anticipated, the lowest emissions solution is the social cost of carbon method. The second lowest total emissions come from the ineffective tax or first pass of the two-pass method, showing that no trade with California actually lowers emissions more than any method. This is due to the high-cost low-carbon resources in California. The highest cost method (social cost of carbon) produces the lowest emissions, while the lowest cost solution (No-GHG) produces the highest emissions. These set up the ranges of possible cost and emissions outcomes, and allow us to examine the solutions in-between. The single pass method has low operating costs (excluding GHG costs), but much higher emissions compared to the two-pass method.

Table 4-8 Emissions per node in tons of CO_2/hour

	No-GHG	Single Pass	Two-Pass		Tax		Social Cost of Carbon
			First Pass	Second Pass	Effective (<20)	Ineffective (>20)	
CA	394.50	396.40	518.20	440.26	431.20	518.20	423.66
NW	175.00	175.00	141.46	175.00	175.00	141.46	175.00
SW	803.93	799.95	626.17	708.00	726.99	626.17	684.40
Total	1,373.43	1,371.35	1,285.83	1,323.26	1,333.19	1,285.83	1,283.06
Reduction relative to No-GHG	0	-2.1	-87.6	-50.2	-40.2	-87.6	-90.4

A regional fuel cost per unit of emissions reduction can be determined by comparing the change between each model and the No-GHG case (difference in total cost, in Table 4-7, divided by total emissions), seen in Table 4-9. Operating costs for the comparison include the marginal fuel cost, excluding any GHG costs. This cost per ton is the incremental cost of removing carbon, or the social cost of removing CO_2. While the magnitude of the average cost in each case is different, the relative changes for California generators are the same, \$48.52/ton.

This value is the sum of the marginal cost of Gen 1 (increases its output) and GHG allocation dual variable, which is difference between the GHG cost and marginal cost of Gen 1 and the marginal cost of Gen 7 ($38-$17.6). When California reduces its emissions, it is only because of moving Gen 1 up and Gen 7 down in every case, which is why the California marginal cost of emissions reduction is constant.

Since the Northwest costs and emissions are the same as the No-GHG case for all cases except when no trade occurs, this means that no average cost can be calculated. Similar to California, the Southwest has the same cost per unit for all cases except the social cost of carbon case. The totals show a similar incremental cost for the solutions, a higher cost for the cases without trade (first pass and ineffective tax), and the lowest cost for the social cost of carbon case. Since both costs and emissions decrease for the Northwest and Southwest, the values are negative.

Table 4-9 Cost increase per metric ton of CO_2 reduction ($/ton, excluding GHG costs) compared to the No-GHG method

	Single Pass	**Two-Pass**		**Tax**		**Social Cost of Carbon**
		First Pass	**Second Pass**	**Effective (<20)**	**Ineffective (>20)**	
CA	48.52	48.52	48.52	48.52	48.52	48.52
NW	no change	-32.40	no change	no change	-32.40	no change
SW	-14.47	-14.47	-14.47	-14.47	-14.47	-5.73
Total	17.50	26.75	16.61	16.62	26.75	8.10

Even though the single pass costs less than the other methods, the social cost of removing CO_2 is around $17/ton for all second-best[32] methods, about double the cost of the social cost of carbon case. This is also the slope of a line connecting the second-best points on Figure 4-2, discussed next, with the no-GHG solution. The social cost minimization case removes more

[32] Second-best methods are methods which cannot impose a direct tax on all carbon emissions, but given the regulatory constraints of the system, can achieve carbon reductions cost-effectively.

CO_2 at a much lower per unit cost, which highlights the inefficiency of partial geographic coverage of carbon laws.

Emissions and non-GHG operating costs are compared side-by-side for each model in Figure 4-2. As might be expected, the No-GHG model is in the upper left corner, with low costs and high emissions. In the lower right corner is the first pass of the two-pass model and the ineffective tax, both with no imports into California. These models have the highest costs, but also consequently have low emissions. Lower emissions for no import cases might not be a trend for systems with different characteristics, especially, if the regulated system has higher emissions rates than surrounding regions (shown briefly in Section 4.4.3). Falling between these cases is the effective tax and the second pass of the two-pass system. Neither system has the lowest emissions or the highest costs.

Figure 4-2 Comparison of total opportunity costs (without carbon costs, $) versus total system emissions (tons)

Meanwhile, the social cost of carbon model deviates from the downward sloping line connecting the other four, with the lowest emissions and relatively low costs. The social cost system dominates all the solutions except No-GHG and single-pass by having both lower costs and lower emissions, again highlighting the large inefficiencies of partial coverage of GHG regulations. Otherwise, no method dominates the other methods. Assuming that the social cost solution is not politically feasible, implementation of another method involves spending more money for more emissions reduction, which is a value judgment. Thus, using the economist's definition of second-best,[33] all the other solutions are efficient ways of achieving the given level of reduction, given the political constraint, and are second-best. This conclusion for second-best methods may be highly dependent on the system in question, and I anticipate that the shape of a graph like Figure 4-2 can change significantly system to system. However, the social cost of carbon method will always dominate the second-best methods, because it directly incorporates the cost of GHG emissions into the dispatch decisions. Future research should attempt to obtain analytical results to assess the generality of these conclusions or investigate other assumptions and systems.

As another analysis of system emissions, the methods can be compared against the No-GHG case. Table 4-10 shows the percent increase of each simulation's emissions compared to the No-GHG method by region and in total. A positive value represents an increase, while a negative value shows a decrease. The single pass is only slightly lower in total emissions, while the social cost of carbon method shows a 7% decrease. The region-by-region percentages show that while some areas reduce emissions, in all cases, California increases emissions compared to the No-GHG case. This demonstrates the value of the EIM; although California plants might output more, the region will benefit from any carbon pricing method. This is also atypical as to what might be expected in general with emissions caps, in which emissions in the regulated region would be expected to decrease, with leakage increasing emissions outside (as, e.g., in RGGI, see [31]). The reason is that, first, imports into California are already at their upper bound. Second, California is much cleaner than one of its neighboring regions (the Southwest), so a regulatory system that motivates reductions in emissions in imported power might shift

[33] The theory of second best states that if the optimal solution cannot be attained due to a constraint, then the next-best optimal solution can only be gained by deviating from that solution (all Paretian conditions) [241].

production from a dirty region (the Southwest) to a clean region (California). The capped imports can be seen in Table 4-4, where the imports (or net exports) from the No-GHG case are higher than all other method's net exports.

Table 4-10 Change in CO_2 emissions due to regulation (percent increase from No-GHG case)

	No-GHG	Single Pass	Two-Pass		Tax		Social Cost of Carbon
			First Pass	Second Pass	Effective (<20)	Ineffective (>20)	
CA	-	0.48%	23.87%	10.39%	8.51%	23.87%	6.88%
NW	-	0%	-23.71%	0%	0%	-23.71%	0%
SW	-	-0.50%	-28.39%	-13.55%	-10.58%	-28.39%	-17.46%
Total	-	-0.15%	-6.81%	-3.79%	-3.02%	-6.81%	-7.04%

4.4.2.d Prices and Profits

Energy prices for each model are shown in Table 4-11, as are CO_2 prices. In general, the energy price at a location is the increase in the objective function (total cost), resulting from a 1 MW increase in demand at that location. Without congestion, all prices would be the same throughout the network. The single-pass, effective tax, and social cost of carbon models do not have any congestion, whereas all others have one line at its maximum capacity. The prices for uncongested models shown in Table 4-11 are the same for regions outside of California, but prices in California reflect GHG costs.

Table 4-11 Energy prices per node and carbon price ($/MWh)

	No-GHG	GHG only in CA	Single Pass	Two-Pass		Tax		Social Cost of Carbon
				First Pass	Second Pass	Effective (<20)	Ineffective (>20)	
CA	33.35	38.00	38.00	16.90	38.00	38.00	38.00	38.00
NW	25.47	27.80	17.60	16.20	17.06	17.60	16.20	38.00
SW	17.60	17.60	17.60	17.60	17.60	17.60	17.60	38.00
CO_2 Price	N/A	N/A	20.4	0	20.67	20.4	21.10	N/A

The energy price calculation for the GHG models is different from the No-GHG and social cost of carbon methods, since it includes the cost of greenhouse gas emissions inside California but not outside. The calculation for each price is shown in Section 4.3, under each model's respective sub-section. To summarize, the second-best methods use the dual variable of either the GHG allocation constraint or the tax constraint as an adder to determine prices in California. Prices outside of California are calculated without the GHG adder. Similar to the price calculation method described by CAISO [148], incorporating the GHG price adder incorporates the GHG cost into California prices and excludes it from prices in the rest of the West. California residents will pay for the GHG costs, while residents in the unregulated states will not. For the single- and two-pass models, the CO_2 price is the dual variable of the GHG allocation constraint in (4-9) and (4-15). For the tax model, it is the dual variable of the California flow constraint in (4-26).

The results can be summarized as follows. The energy price for the social cost of carbon is reflective of Gen 1's fuel and GHG cost, which is marginal throughout the uncongested network. The single-pass is also uncongested, but the price outside California is lower because it does not include the GHG dual variable, which is 20.4 $/MWh. This variable, the CO_2 price for both the single-pass and tax, is the difference between the marginal and GHG cost of Gen 1 ($38/MWh) and the marginal cost of Gen 7 ($17.60). Compared to the No-GHG model, these two generators increase and decrease output respectively. The two-pass method produces some congestion, but still shows similar pricing compared to the single-pass method. The first pass of the two-pass method is not used for settlements, but is shown as a reference without the GHG constraint dual adder.

The prices resulting from the effective tax method are the same as the single-pass method, and almost the same as the two-pass method. The dual variable on the import/export constraint in the tax method (constraint (4-26)) has the same value as the GHG export constraint in the single pass method (constraint (4-9)), 20.4 $/MWh. This is also the value at which the tax changes from effective to ineffective. This value must be included in pricing for the tax method, else the prices will not be supporting, meaning the price will be too low to incentive the generator to produce (the generator would operate at a loss).

Profits for each generator based on the prices in Table 4-11 are shown in Table 4-12 and Table 4-13. The profits in Table 4-12 are calculated assuming that allowances are allocated for free to the generators, i.e., the costs exclude the GHG costs. On the other hand, if California generators are required to purchase allowances, the profits would be those found in Table 4-13, where GHG costs are included. If all generators in the West are required to purchase allowances, profits would be negative in the Northwest and Southwest. Prices outside California do not include the GHG adder, which would mean prices are not supporting outside California. Generators in the Northwest and Southwest would require another means of cost recovery if required to purchase allowances, or pricing would need to be adjusted.

For this example, profits are highest in the social cost of carbon model, since the price of power on average goes up by more than the average emissions rate times the social cost. Other systems might have different results, for instance if the marginal generators are clean generators and the dirty generators are at capacity, so that the average emissions rate is greater than the marginal emissions rate. Taxing emissions might increase or decrease profits, depending on the average emissions rate and the marginal emissions rate [182], [183]. The GHG only in CA and No-GHG models have the next highest profits under both purchased and free allowances, due to congestion in the network leading to higher profits in the Northwest. Because all generators in the Northwest are at maximum capacity, delivering an additional MW to the Northwest requires increased 0.5 MW increased output from generators in both California and the Southwest.

The single-pass and effective tax methods produce the same prices in all locations because there is no congestion. If allowances are purchased (Table 4-13), the total profits are the same. Gen 1 and Gen 7 are dispatched to different quantities in the simulation, but both are marginal generators and do not make a profit (the price is equal to their marginal costs). However, if allowances are allocated for free (Table 4-12), the change of dispatch level between the two methods returns higher profits for the effective tax case compared to the single-pass. Without the ability to sell power to California, the ineffective tax produces the lowest profits when allowances are purchased. In that case, Gen 5 is marginal in the Northwest, where in the other methods it is producing at full capacity.

Profits are understandably impacted by whether or not the allowances are allocated to each generator for free or are purchased. If the generators are purchased, profits are reduced significantly in this example. The last row of Table 4-13 shows the decrease in profits for California generators when they are required to pay for allowances. The marginal generator breaks even in this case, and because the other generators are at maximum capacity, they produce the same profits for each model. If the allowances are allocated for free, and prices still include the GHG adder, all generators will make additional profits. Rather than pay the generators the additional cost of allowances, a system operator might choose to return the surplus to California consumers who are paying a higher price for electricity. Alternatively, the state can choose to use this surplus for carbon reduction activities. The means of income distribution is necessary area for further research.

Table 4-12 Profits per generator assuming GHG are allocated for free ($)

	Region	No-GHG	GHG only in CA	Single Pass	Two-Pass		Tax		Social Cost of Carbon
					First Pass	Second Pass	Effective (<20)	Ineffective (>20)	
Gen 1	CA	139	264	296	N/A	1041	887	2,366	759
Gen 2	CA	1,378	2,308	2,308	N/A	2308	2,308	2,308	2,308
Gen 3	CA	3,038	5,130	5,130	N/A	5130	5,130	5,130	5,130
Gen 4	NW	1,493	1,842	312	N/A	230	312	102	3,372
Gen 5	NW	1,855	2,320	280	N/A	171	280	0	4,360
Gen 6	NW	5,095	5,560	3,520	N/A	3,411	3,520	3,240	7,600
Gen 7	SW	0	0	0	N/A	0	0	0	-
Gen 8	SW	384	384	384	N/A	384	384	384	8,544
Gen 9	SW	-	-	-	N/A	-	-	-	-
Gen 10	SW	-	-	-	N/A	-	-	-	3,880
Total		13,379	17,808	12,230	N/A	12,678	12,821	13,530	35,953

Table 4-13 Profits per generator in California assuming allowances are purchased ($)

	Region	GHG only in CA	Single Pass	Two-Pass First Pass	Two-Pass Second Pass	Tax Effective (<20)	Tax Ineffective (>20)	Social Cost of Carbon
Gen 1	CA	0	0	N/A	0	0	0	0
Gen 2	CA	456	456	N/A	456	456	456	456
Gen 3	CA	540	540	N/A	540	540	540	540
Total including profits from NW and SW		9,821	5,492		5,195	5,492	4,722	28,752
Decrease in CA Profits from Table 4-12		6,706	6,738	8,808	7,483	7,329	8,808	7,201

4.4.2.e Carbon Leakage and Contract Shuffling

Other than the No-GHG method, each model changes the dispatch in an attempt to lower emissions in the overall system, based on the incentives embodied in each model. The goal of these methods is to reduce overall emissions, including the impact of leakage and contract shuffling, meaning that low cost imports into California should not also increase emissions elsewhere in the network. I consider two definitions of leakage here:

1) *The traditional definition is the increase in emissions outside of the regulated region*. This calculated simply by comparing the rest of the West (ROW, the sum of Northwest and Southwest) emissions under the regulated regime with a no-regulation (No-GHG) baseline. Leakage can then be expressed as a fraction: (increase in nonregulated region emissions)/(decrease in regulated region emissions). In much of the literature, this is the common definition of carbon leakage [28]–[31], [172]. It is most relevant to cap-and-trade systems that only apply to the regulated region and do not attempt to attribute emissions to imports or regulate them (e.g., RGGI, Clean Air Act SO_2 Trading Program, and European ETS). Leakage using this definition can occur if energy imports from the non-regulated area to the regulated area increase and if the regulations provide no incentive or ineffective incentives to limit measures in the non-regulated area.

2) *The alternative definition proposed in this chapter is most applicable when a regulatory system attempts to quantify and regulate emissions associated with imports from non-regulated systems* (as California does). The alterative definition of leakage calculates the difference between (a) the accounted for reduction of emissions (including emissions within the regulated region, and emissions associated with imported power from the nonregulated region that are subject to the regulatory system) relative to a baseline, and (b) the actual total reduction of emissions in both the regulated region and the nonregulated region relative to the actual reduction. If this difference is divided by (a), this is a percentage measure of leakage. Definition (2) will be the same as (1) if the accounted for emissions in imports into the regulated region equals the actual difference between emissions outside the regulated region under the model in question, and those emissions if instead imports are restricted to be zero. This is a measure of how accurate the accounted for reductions under the regulatory scheme (including any that are associated with imports) are relative to the actual reductions totaled across the regions. It is of interest when a regulatory system, such as California's, attempts to limit leakage and/or contract shuffling by counting emissions associated with imports and assigning carbon costs to them (through taxes or required purchases of emissions allowances).

Calculations based on the first definition of leakage, i.e., increase in emissions outside the regulated region compared to decrease in emissions in the regulate region, are shown in Table 4-14. Given this definition, the example presented in this chapter has no leakage, or has reverse leakage. Emissions in the Northwest and Southwest are highest in the No-GHG case, and lower in all models that consider emissions. This is a surprising result, one that results from the specific assumptions and input data used for the example. In particular, there is a constrained transmission line between the Southwest and California, which does not allow additional imports into California. While overall emissions decrease due to redispatch of existing resources, emissions inside California increase. This type of result would not be typical in real-world systems, since there are more interconnections between two regions than shown here. Table 4-14 shows the emissions in the rest of the West decrease compared to the No-GHG case, and California emissions increase. Thus, examining leakage using this definition does not produce a meaningful result. Since the change in the rest of the West emissions is negative and it is divided by a large increase of emissions in California, this definition of leakage yields a misleadingly positive and large result.

Table 4-14 Calculation of leakage from definition (1) above (tons)

	No-GHG	**Single Pass**	**Two-Pass**		**Tax**		**Social Cost of Carbon**
			First Pass	**Second Pass**	**Effective (<20)**	**Ineffective (>20)**	
Emissions in ROW	978.93	974.95	767.63	883.00	901.99	767.63	859.40
Change in ROW emissions (ΔROW)	N/A	-3.9	-211.3	-95.9	-76.9	-211.3	-119.4
Emissions in CA	394.50	396.40	518.20	440.26	431.20	518.20	423.66
Change in CA emissions (ΔCA)	N/A	1.9	123.7	45.8	36.7	123.7	29.2
Leakage Definition (1) (-ΔROW/ΔCA)	N/A	205%	171%	209%	210%	171%	409%

By using additional data from the simulation outcomes, we can calculate leakage using the second definition. This calculation incorporates the regulation's errors in accounting for emissions in imports in cases where the cap-and-trade systems attempts to quantify and penalize emissions associated with imports, as California attempts to do. The second method to calculate leakage compares two elements, labeled (a) and (b) above.

Part (a) compares the accounted for reduction in emissions relative to an adjusted No-GHG baseline. The baseline uses California emissions from the No-GHG case, and the difference between the rest of the West emissions in the No-GHG case and the first pass. The baseline examines the difference between the total emissions from the No-GHG case and emissions in the rest of the West only due to their own needs (which is the definition of results from the first pass). The baseline is then compared to each model's specific emissions in California and the emissions due to regulated import. The import is calculated differently depending on the model.

For the single- and two-pass method, the imports are calculated with the greenhouse gas variable introduced in Section 4.3.2, p_{gt}^{GHG}, multiplied by the plant's emissions rate. For the tax cases, a similar value can be calculated. Using the lower bound of p_{gt}^{GHG}, shown in constraint (4-9), imports can be calculated by taking the total imports (total supply less demand in the Northwest and Southwest), and choosing p_{gt}^{GHG} from the lowest cost greenhouse gas bids. The simple method to calculate this value for the tax and social cost of carbon case is shown in (4-27)-(4-29). The calculated results are shown in Table 4-15 as "GHG Emissions Imports as Accounted for by Regulation (RegI)," and the total values for part (a) are shown as ΔRegI.

$$\min \sum_{\forall t} \sum_{\forall g \notin CA} C_g^{GHG} p_{g,t}^{GHG} \tag{4-27}$$

$$\sum_{g \notin CA} p_{g,t} - \sum\nolimits_{i \notin CA} D_{i,t}^{\max} = \sum_{g \notin CA} p_{g,t}^{GHG} \qquad \forall g \notin CA, t \tag{4-28}$$

$$p_{g,t}^{GHG} \leq G_{g,t}^{GHG} \qquad \forall g \notin CA, t \tag{4-29}$$

Part (b) is compares the change in total emissions from the No-GHG case to the model specific emissions, called ΔTot. This value is the actual reduction in total emissions, whereas Part (a) shows the reduction due to accounted for emissions. The values for each part are shown in Table 4-15, based on the relationships below.

- ΔRegI = (Baseline) – (Accounted for Emissions)

  = [(*No-GHG* CA)+((*No-GHG* CA)–(*First Pass* ROW))]–[(*Model* CA)+(*Model* GHG imports)]
- ΔTot = (*No-GHG* Total) – (*Model* Total)
- Leakage = 1 – ΔTot/ΔRegI

Leakage using the second definition is shown in gray in Table 4-15. This definition of leakage extends beyond most found in the literature to include the impacts accounting for regulated emissions. As was evident from the leakage in Table 4-14, this case might not demonstrate the traditional definition of leakage, but still shows the impact of shifting imports between generators inside and outside California. Leakage is zero when there are no imports into California, seen in the first pass and the ineffective tax. The single pass has 98.5% leakage; there is only 2.1 tons of actual emissions reduction (ΔTot), but almost 140 tons accounted for due to regulation. Plants that would have been used to serve local demand are now being counted for California demand, and other dirtier plans are filling the difference. A similar phenomenon happens with the effective tax. Supply in California is being fulfilled by Southwest supply, but to a lesser degree, resulting in 70.1% leakage. In comparison, the two-pass method increases low emitting supply in California and reduces imports. The total emissions reduce more than the accounted for emissions from regulation and the net result is a negative percentage for leakage.

Table 4-15 Alternate calculation of leakage (tons/hr)

	No-GHG	**Single Pass**	**Two-Pass**		**Tax**		**Social Cost of Carbon**
			First Pass	**Second Pass**	**Effective (<20)**	**Ineffective (>20)**	
CA Emissions (CA)	394.50	396.40	518.20	440.26	431.20	518.20	423.66
ROW Emissions (ROW)	978.93	974.95	767.63	883.00	901.99	767.63	859.40
Total Emissions	1373.4	1371.4	1285.8	1323.3	1333.2	1285.8	1283.1
Actual Reduction in Total (ΔTot)	N/A	2.1	87.6	50.2	40.2	87.6	90.4
GHG Emissions Imports as Accounted for by Regulation (RegI)	1258	70.0	0.0	134.8	40.0	0.0	46.5
Δ GHG as Accounted for Regulation (ΔRegI) [Baseline* – (CA+RegI)]	N/A	139.40	87.60	30.78	134.60	87.60	135.64
Leakage Definition (2) [1-ΔTot/ΔRegI]	N/A	98.5%	0.0%	-63.0%	70.1%	0.0%	33.4%
Change in imported emissions compared to No-GHG (ΔImport)	N/A	1188	1258	1123	1218	1258	1212
Contract Shuffling [1-ΔROW/ΔImport]	N/A	100%	117%	109%	106%	117%	110%

*Baseline
= California Emissions from No-GHG+(ROW Emissions from No-GHG – First Pass ROW Emissions)
=1373.4-767.63 = 605.8

Similar to the calculations from [11], the total contract shuffling can be calculated using the change in the rest of the West emissions and the apparent change in import emissions into California. The change in ROW emissions is shown in Table 4-14, and the change in imports is shown in Table 4-15 as (ΔImport). The amount of contract shuffling is over 100% in all GHG models. Since no additional power can be imported into California, all regulated emissions decreases are due to shifting contracts among sources of imports from the rest of the West. In this case, contract shuffling was utilized to reduce emissions. However, unlike the results in this example, contract shuffling might not always work to the advantage of the regulated region, especially if the contracts are outside the market. For instance, if a generator not dispatched by the ISO is able to sell to an unregulated state.

Leakage and contract shuffling are both complicated concepts, each with different definitions in literature and few papers which attempt to quantify their impact. One or both can result when neighboring regions have asymmetrical policies in place, both for carbon emissions and other greenhouse gases. The effect of leakage will vary depending on the accounting method used for imports. Neither leakage nor contract shuffling have necessarily negative impacts; if overall emissions decrease, then shifting emitting plants from one regions to another has a net benefit for the system. However, this does not account for local emissions or effects on communities. While outside the scope of this chapter, it can be a criterion for method evaluation, discussed briefly in Section 4.5.2.

4.4.3 Modified Network

Although emissions reduced using the GHG models in the previous example, emissions in the rest of the West also decreased. This is an unusual result in the sense that trade is limited in the No-GHG case, making results with and without GHG costs in California identical. This is due to the restricted line flow between the Southwest and California. The total amount of imports could not extend beyond the No-GHG case, which is why it did not change total dispatch when including GHG costs for California plants. That case shows the impact of a congested network and the benefits of an alternative definition of leakage. By modifying the network and generator costs in California, a network can be examined in which leakage can occur through increases in imports, which is the more typical concern from environmental groups (for instance, as articulated by the Sierra Club [184]). The original generator and network information in Table 4-1 and Table 4-2 is modified with shaded text in Table 4-16 and Table 4-17. Here, generators in California have comparable marginal fuel costs to the rest of the West, and higher emissions rates compared the previous example (now comparable to the Southwest). This example simulates a situation in which a capped markets; units have emissions rates similar to neighboring uncapped regions.

The same five models are simulated, and notably, there are now two distinct No-GHG cases. Without any carbon price, the emissions in the entire region are high. With GHG prices only in California, emissions decrease overall, but increase in the rest of the West, resulting in leakage by the first definition in the previous section.

Table 4-16 Generator characteristics

	Location	Marginal Cost ($/MWh)	Capacity (MW)	CO_2 Rate (kg/MWh)	GHG bid (MW)	GHG cost ($/MWh)
Gen 1	CA	15.2	250	1429	-	24.3*
Gen 2	CA	15.9	200	1342	-	22.82*
Gen 3	CA	16.0	450	929.4	-	15.8*
Gen 4	NW	15.52	150	500	150	8.5
Gen 5	NW	16.2	200	500	200	8.5
Gen 6	NW	0	200	0	180	0
Gen 7	SW	17.6	400	1216	400	20.67
Gen 8	SW	16.64	400	1249	400	21.23
Gen 9	SW	19.4	450	1171	450	19.91
Gen 10	SW	18.6	200	924	200	15.71

*These values are not used as GHG cost bids in simulations, rather are added to the marginal costs and bid as energy costs. The resulting California marginal cost bids total 39.5, 38.72, 31.80 $/MWh respectively. Only the No-GHG case uses the marginal cost values from the table.

Table 4-17 Network characteristics

Line/Node			CA	NW	SW
	Max Capacity	Load	890 MW	303 MW	684 MW
Line 1	255 MW	PTDF Values	+0.3333	–0.3333	0
Line 2	120 MW		+0.3333	+0.6667	0
Line 3	350 MW		–0.6667	–0.3333	0

Using the same two definitions of leakage from Section 4.4.2.e, Table 4-18 compares the amount of leakage in a system with larger transmission capacity, allowing increased imports into California. The first definition of leakage is the change in emissions from the unregulated region (ROW), which can be expressed as a ratio to the change of emissions within the regulated region (CA). The comparisons are between the No-GHG case and all remaining models, including the case with GHG prices in California alone. Calculations using the second leakage definition (described in Section 4.4.2.e) are shown in the last row of Table 4-18. As a reminder, leakage using definition (2) is the difference between the actual reduction in the West's emissions and the reduction as accounted for by the regulation (including imports). Both leakage results are highlighted in gray.

Table 4-18 Alternate calculation of leakage for modified system (tons/hr)

	No-GHG	**GHG only in CA**	**Single Pass**	**Two-Pass**		**Tax**		**Social Cost of Carbon**
				First Pass	**Second Pass**	**Effective (<26)**	**Ineffective (>26)**	
Emissions in CA	1043.7	264.8	592.5	1029.4	686.5	264.8	1029.4	418.1
Change in CA emissions (ΔCA)	N/A	-778.9	-451.2	-14.3	-357.2	-778.9	-14.3	-625.6
Emissions in ROW	759	1452.8	1096.6	767.6	1006.9	1452.8	767.6	1196.2
Change in ROW emissions (ΔROW)	N/A	693.8	337.6	8.6	247.9	693.8	8.6	437.2
Total Emissions	1802.7	1717.6	1689.1	1797	1693.4	1717.6	1797	1614.3
Change in Total Emissions (ΔTot)	N/A	-85.1	-113.6	-5.7	-109.3	-85.1	-5.7	-188.4
Leakage Definition (1) (-ΔROW/ΔCA)	N/A	89%	75%	60%	69%	89%	60%	70%
GHG Emissions Imports as Accounted for by Regulation (RegI)	0	552.3	191.6	0	258.7	552.3	0	345.2
Δ GHG as Accounted for Regulation (ΔReg) [Baseline*–(CA+RegI)]	N/A	218	251	5.7	89.9	218	5.7	271.8
Leakage Definition (2) [1-ΔTot/ΔReg]	N/A	61%	55%	0%	-22%	61%	0%	31%

*Baseline
= California Emissions from No-GHG + (ROW Emissions from No-GHG – First Pass ROW Emissions) = 1035

This example shows a case where leakage occurs in the nonregulated region (ROW) due to shifting emissions from California to the West. The new values for emissions in California create more leakage that might occur in actual system operations. However, other regions in the Eastern U.S. have higher emitting resources in-state and lower emitting resources out of state. In this case, leakage might be a significant issue. Further analysis of this and other similar cases are necessary to evaluate the impact of leakage between the different second-best or GHG models.

4.4.4 Sensitivity Analysis for the Carbon Price

In order to test the sensitivity of single-pass, two-pass, and social cost of carbon models, different carbon prices were tested for all the models. The following tables show the results of the models with carbon prices of $10/ton, $15/ton, and $20/ton. The base case, which resulted in the previous section's results, is $17/ton. The results for a price of $10/ton or $15/ton are the same, and therefore only shown once. A $25/ton price was also tested, with the only change being to the social cost of carbon model, therefore it is shown in the emissions tables as a separate column. A positive value indicates an increase in costs and an increase in emissions. Likewise, a negative value indicates a decrease in costs and decrease in emissions.

With a lower carbon price, total costs decrease for all models, although costs increase for the Southwest. Even though costs are lower, the total emissions increase. This outcome follows common logic, if carbon is not valued as highly, emissions will increase. The reverse is true when the carbon price increases. With a price of $20/ton, prices increase and emissions decrease. At an even higher price, $25/ton, only the social cost of carbon model shows a change dispatch. The costs increase even higher, but the emissions decrease by just over 58 metric tons compared to the $17/ton price.

The social cost of carbon method is the only model that changed when faced with a higher carbon price, which confirms the dominance of the social cost of carbon method over the second-best models. The single- and two-pass methods both respond to a lower cost per ton of carbon, but only the two-pass method reduces emissions when the cost increases. For this example, this would suggest the two-pass method is better able to respond to a changing carbon price, dominating the single-pass.

Table 4-19 Change in operating costs per node in \$/hr, (CO_2 = 10 or 15 \$/ton) – (Base Case)

	Single Pass	**Two-Pass**		**Social Cost of Carbon**
		First Pass	**Second Pass**	
CA	(124)	-	(2,998)	(1,910)
NW	-	-	-	-
SW	58	-	1,520	885
Total	(67)	-	(1,477)	(1,026)

Table 4-20 Change in operating costs per node in \$/hr, (CO_2 = 20 \$/ton) – (Base Case)

	Single Pass	**Two-Pass**		**Social Cost of Carbon**	**Social Cost of Carbon (25 \$/ton)**
		First Pass	**Second Pass**		
CA	-	-	3,586	1,088	6,201
NW	-	-	(763)	-	(1,088)
SW	-	-	(832)	(476)	(1,597)
Total	-	-	1,991	611	3,515

Table 4-21 Change in emissions per node in kg/hr, (CO_2 = 10 or 15 \$/ton) – (Base Case)

	Single Pass	**Two-Pass**		**Social Cost of Carbon**
		First Pass	**Second Pass**	
CA	(1,897)	-	(45,755)	(29,157)
NW	-	-	-	-
SW	3,976	-	92,631	61,128
Total	2,080	-	46,876	31,972

Table 4-22 Change in emissions per node in kg/hr, (CO_2 = 20 \$/ton) – (Base Case)

	Single Pass	**Two-Pass**		**Social Cost of Carbon**	**Social Cost of Carbon (25 \$/ton)**
		First Pass	**Second Pass**		
CA	-	-	54,741	16,599	94,644
NW	-	-	(23,560)	-	(33,590)
SW	-	-	(57,470)	(35,745)	(119,904)
Total	-	-	(26,289)	(19,146)	(58,850)

4.5 DISCUSSION

A region that values an externality more than its neighbors faces challenges when both regions engage in a common market. This question has come to the forefront with the implementation of California carbon emissions trading and concerns over carbon leakage. It can generally arise between any two or more countries with different pollution trading or penalty mechanisms. In the case examined in this chapter, California established a cap-and-trade system; it values reducing carbon emissions from power plants and imported power. The rest of the West has yet to implement a similar scheme, which has resulted in stakeholders being concerned that expansion of power markets in the West will result in increased imports of coal-based power to California and will weaken the effectiveness of California's emissions limits [184], [185]. The Western grid is well connected, and a real-time integrated market operates every 5 minutes with many Western utilities and the CAISO. California is often dependent on surrounding states for power, and similarly would prefer to export renewable power if it has excess. However, exporting or importing power while effectively reducing emissions can be a challenge.

Although it is the first best option, the social cost of carbon method cannot be implemented in the West given current state policies. Residents and legislators in neighboring regions would need to opt to dispatch their own system based on greenhouse gas costs, and similarly pay for the allowances (either generators would be required to pay for allowances or customers would be required to pay a premium for electricity). Until neighboring states enact a greenhouse gas policy, second best options must suffice. At present, British Columbia has a carbon policy [186], and other Western states have proposed [187] or failed to pass [188] carbon legislation. However, enacting legislation in some but not all neighboring regions will also complicate the dispatch of the system, since each region would require its own constraints and variables.

4.5.1 Results Summary

By imposing a carbon price (in the form of an allowance requirement) upon generation within one jurisdiction that exists within a larger electricity market region, leakage and contract shuffling can occur. In other words, the overall emissions reduction in the entire region may be less than the apparent emissions reduction in the target jurisdiction. There are several existing proposals to manage trade and leakage from centralized power dispatch, each with benefits and downsides. Three main proposals are compared against simulations without an emissions price (No-GHG) and with a system-wide emissions price (social cost of carbon) in Section 4.2.14.4.2. A summary of results can be found in Table 4-23.

Table 4-23 Summary of key results

	No-GHG	Single Pass	Two-Pass		Tax		Social Cost of Carbon
			First Pass	Second Pass	Effective (<20)	Ineffective (>20)	
Total Operating Costs ($/hr, other than CO_2 charges)	$34,642	$34,677	$36,983	$35,472	$35,310	$36,983	$35,373
Total Emissions (tons/hr)	1373.4	1371.4	1285.8	1323.3	1333.2	1285.8	1283.1
Total Social Cost ($/hr, including socialized CO_2 costs)	$57,990	$57,989	$58,843	$57,969	$57,973	$58,843	$57,184
Cost to CA Consumers ($/hr)	$29,682	$33,820	-	$33,820	$33,820	$33,820	$33,820
Incremental Cost to CA Consumers* ($/ton)	N/A	-2,069	-	-83	-103	-47	-46
Incremental Total Cost** ($/ton)	N/A	-17.5	-	-17	-17	-27	-8.1
Leakage Definition (2)***	N/A	98.5%	0.0%	-63.0%	70.1%	0.0%	33.4%

*Incremental Cost to California Consumers = (ΔCA Cost/ΔTotal Emissions)

**Incremental Total Cost = (ΔOperating Cost with GHG Costs/ΔTotal Emissions)

***Leakage Definition (2) = 1 – ΔTot/ΔRegI, defined in 4.4.2.e where RegI is the accounted for regulated emissions

I hypothesized that the lowest cost and highest emissions solution would be the No-GHG case, while the lowest emissions solution would be the social cost of carbon case. The remaining three GHG proposals would then fall somewhere between the two extremes. The results confirm this hypothesis and show each of the three proposals increase costs beyond the No-GHG simulation and also yield higher emissions than the social cost of carbon solution (in which a single price is applied to all emissions throughout the West), as shown in rows 1 and 2 of Table 4-23. The first best option, the social cost of carbon case, produces the lowest emissions, but that does not imply it will also be the highest cost simulation. If carbon costs are considered for all plants (assuming allowances must be purchased), the social cost of carbon case is actually the lowest cost solution, seen in row 3, and its fuel costs are about the same as the ineffective tax solutions. But if only California plants must purchase emissions allowances, then the social cost of carbon model produces the highest cost solution. Whether or not these allowances are purchased or provided will have a significant impact on generator profits, and should be considered when analyzing model formulations.

The example in this chapter is based on California issues, but is not a direct representation of the California system. The conclusions drawn are representative of one type of constrained system, and not intended to directly suggest policy implications of AB32. The system shown in this chapter is an extreme case examining how emissions can be accounted for when imports are constrained. Although California imports are not generally constrained, it might apply to other networks or future bottlenecks. Future work can assess different system states. For instance, other simulations can include a case where imports are not constrained due to congestion, a case with different load levels throughout the network, or different types of supply (such as renewable energy). A case that will be of special interest to California would examine the impacts of the duck curve, or the high penetration of solar energy midday that reduces net demand resulting in a large ramp event in the early evening hours. Using a variety of systems states will allow for a holistic assessment of impacts to both the regulated and unregulated system.

If a system operator prioritizes lowering operating costs, the simulations indicate the single-pass method would be the best alternative. However, it is also the method that produces the highest emissions (that includes carbon pricing); very little emissions reductions take place. The single-pass method also has the highest carbon leakage, 98.5%, since it does not encourage emissions reductions in the network, even though the regulatory system's account appears to take credits for large reductions in imported carbon. The two-pass method is better able to capture emissions associated with imports by simulating the market without California and adjusting based on the first pass. The resulting second pass reduces emissions compared to the single-pass while increasing costs, at an incremental cost per ton of about twice what can be achieved by an efficient West-wide system (social cost of carbon method). It has negative leakage, because there are more actual reductions than accounting reductions, i.e., California is counting a reduction of 30 tons, whereas the total reduction was 50 tons. This might be problematic when assessing the efficacy of the emissions reductions.

The tax model results in two separate cases, one in which the tax limits imports and one in which the tax eliminates imports (does not allow trade between California and the West). The effective tax (one that limits imports) results in total costs and emissions similar to the two-pass method. Although not every network will show such similar results for the two models, the similarities show the tax might be able to reduce system-wide emissions. The ineffective tax (one that eliminates imports) stops all trade between the regions. Emissions are low due to low emitting resources in California, but operating costs and consumer payments are high. Although the magnitude will change, a tax that eliminates imports makes the West-wide market (EIM) worthless.

The optimal tax for a particular system will be different depending on imports, and can vary between neighboring systems. For instance, the tax on imports from the Northwest might be different than a tax on imports from the Southwest because congestion results in generators in both regions being on the margin ("basic", in the terminology of optimization). If implementing a tax model, the regulated region can assess if the taxes should be adjusted depending on location. In theory, the tax could reflect the marginal emissions rate in the surrounding region, which can differ due to congestion in the network. In the example in this chapter, the optimal rate was the difference between the marginal and greenhouse gas costs in California and the marginal cost of the marginal resource outside of California. Optimal taxes will likely also need to be reassessed and adjusted over time.

Turning to price impacts for the West-wide test case, prices are very similar, but the total cost to California customers varies due to imports. The No-GHG case costs California consumers the least, while the social cost of carbon model and ineffective tax cost the most. Both the social cost of carbon and ineffective tax cost $38/MWh, which is the cost of the marginal generator in California. Among the GHG models, the single-pass method costs the least for California consumers, because California imports cheaper power from both the Northwest and the Southwest. An effective hurdle rate or tax produces similar emissions and costs compared to the two-pass method, and costs the California consumers the least slightly less than the two-pass method. Consumers might prefer the tax to the two-pass; however, if the tax is too high, consumers must pay for high cost in-state resources.

As discussed in Section 4.2.3, many in industry believe the two-pass method presents incorrect incentives for generators. Due to the two passes, a generator might construct an offer that avoids dispatch in the first pass but ensures dispatch in the second. This might encourage dishonest bidding, which is not modeled in this chapter. Because this analysis assumes truthful bidding, future work can examine the impact of dishonest bidding in each case; emissions and costs might differ greatly if generators are inflating their bids. These concerns led the CAISO to modify their proposal further, eliminating the two-pass method. While it will not be implemented in a market, the two-pass solutions can lead to certain insights. For modeling purposes, the first pass establishes base generation needed in each region, before trade occurs. Identifying resources that increase or decrease output due to trade is difficult without a variable identifying that information. However, this model may still be useful as a baseline for future research that explicitly calculates resource shuffling, since the resources that respond to trade change output.

4.5.2 Policy Recommendations

None of the three second-best regulatory methods dominate the other in that none can both lower costs and lower emissions more than any other method. This emphasizes the difficulty in finding the second-best method for pricing carbon, assuming that the first best method of implementing a social cost throughout the West is not politically feasible. The single-pass method might not reduce emissions to the same extent as the other methods, but it is simple to implement. Each resource bids its allowance cost, and one linear model is used. As discussed in Section 4.2.3, the two-pass method identifies the carbon dispatch due to California, but might have poor long-term impacts on bidding strategies. The tax reduces emissions, but is difficult to determine. A tax that is too high will stop all trade while a lower tax has higher leakage compared to the two-pass method. No one method is without fault, a point emphasized on a Feb. 22nd, 2018 CAISO stakeholder call.[34] In comparing the methods there is a conflict between appropriate prices, low costs, and reducing emissions.

It should also be noted that the three methods plateaued in response to higher carbon prices, i.e., the dispatch did not change as the price increased. Only the social cost of carbon method continued to adjust dispatch in response to higher carbon prices. Although the carbon price has been low in cap-and-trade auctions, allowance prices might continue to rise as renewable penetration increases, as was implemented in British Columbia. In choosing a method, the operator will need to prioritize characteristics of the method that are most important to stakeholders. There might not be a single second-best, or a single dominant method that can reduce carbon emissions cost-effectively subject to the constraint that California cannot impose a tax or shadow price on all carbon emissions by power plants in other states.

In assessing different carbon methods, there are several criteria decisions makers can use to determine which second-best method is best suited for their needs. A list of possible criteria is

[34] Presentation available,
www.caiso.com/Documents/Agenda-Presentation-RegionalIntegration-EIMGreenhouseGasCompliance-Feb22_2018.pdf

given in **Error! Reference source not found.**. The first two criteria are the axes of Figure 4-2: total operating costs and emissions. These are the criteria that indicate whether a policy is efficient (has the lowest cost of achieving the target), the focus for most economists.

Table 4-24 Criteria that can be used to assess GHG models

Category	**Sub-categories**
Operating Costs	Including only marginal costs Including greenhouse gas costs (allowances)
Emissions	Greenhouse gas emissions Local air pollution Leakage & contract shuffling
Economic efficiency	Changes in consumer and producer surplus Opportunities for gaming, exercising market power
Prices	What costs should be incorporated in prices Ensuring prices are supporting
Political impact	Region/state viewed as green leader Legal concerns (interstate commerce) Environmental justice
Logistics	Ease of implementation (administrative, legislative, software) Transparency

Operating costs must be assessed for all parties involved, in both the regulated region and the unregulated region. However, the regulated region will also need to assess the social outcome: if all regions include a greenhouse gas cost, will the total social cost also decrease? Since the impact of carbon emissions is regional, total reduction in emissions can be more important than a local reduction in emissions. While not directly related to carbon emissions, local air pollution can be estimated along with greenhouse gas emissions. Increasing harm to public health is likely to primary concern if power plants are located in populated areas.

Another aspect of emissions that is crucial to consider is how emissions are counted, both in assessing leakage and contract shuffling [147]. California does not want to increase carbon emissions in outside regions due to their own carbon regulations. Similarly, replacing slightly cleaner imports with high emitting resources outside of the EIM would also counter carbon reduction goals. Assessing any new carbon model should calculate leakage using both the traditional and proposed definitions.

The third type of criterion is distribution of benefits and costs. Any new model should ask who wins and who loses (among the regions and between producers and consumers), calculating gains or losses in consumer surplus and producer surplus [146], [189]. It should further analyze the opportunities for gaming or strategic behavior, and any increased opportunities for a participant to exercise market power. Related, a fourth category for assessment is pricing [143]. Each method must ensure that resulting prices are supporting for both regulated and unregulated participants [27]. It must also determine how prices will be calculated. What part of the GHG costs will be incorporated and who will pay for the additional costs?

The last two criteria are political impact and logistics. Neither necessarily produce quantitative results, but are necessary for actual implementation. Whether or not a region implements a cap-and-trade scheme or a carbon tax is highly dependent on the political will of the legislators and voters [160], [190]. If a region would like to be seen as a green leader, or even an independent state leader, implementing GHG models might be easier. There can be legal challenges, such as concerns about the Commerce Clause of the US Constitution [191], or a push for environmental justice. Logistically, ease of implementation is a primary concern. What are the costs to campaign for the change, to educate participants and consumers, and to implement it in software? Finally, it is more likely to take effect if the process and outcomes are transparent.

Trade and coordination are a challenging task for system operators, one made all the more difficult due to renewable energy integration and emissions targets that are not coordinated among sub-regions in an electricity market. This chapter modeled and simulated electricity and emissions markets that directly incorporate emissions allowances and attempt to reduce carbon leakage. Although no one method dominates the rest, the framework used and criteria discussed can be applied to investigate new methods as they are proposed. The results from this simple example show that the system state can have a significant impact on calculations of leakage. Any future comparisons of different proposed models will benefit from analysis of a range of conditions within the network. Given interest in cap-and-trade schemes across the U.S., further analyses will be necessary.

CHAPTER 5
COORDINATION OF TRADE BETWEEN BALANCING AREAS

5.1 INTRODUCTION

Managing the flow of electricity between adjacent electric grids is a challenging task for any system operator. Difficulties can arise when determining rules for trade, ensuring each side see benefits, and coalescing contradicting policies. These difficulties are inflated when renewable energy and environmental concerns are involved. Wind and solar energy are prevalent in different parts of the country, and not necessarily co-located with load centers. Trading that power across regional boundaries can create problems for system operators, who aim to maximize market surplus given local policies and rules on emissions reductions.

This chapter focuses on two aspects of trade and coordination between regions: trade between time periods and trade considering externalities. Within a single market there are many difficulties to modeling trade, difficulties that are exacerbated by the existence of multiple market auctions in time. Markets in the U.S. operate a day-ahead market cleared on an hourly basis and a real-time market cleared on a five and/or fifteen minute basis. As discussed in previous chapters, electricity is sold in both day-ahead markets and real-time or balancing markets. Trade between regions is not isolated to a single market, the mechanisms of each market impact overall market efficiency.

In order to accommodate the variability and uncertainty from renewable energy, many in the literature have proposed changes to operations, including increasing the size of balancing areas (BAs). Balancing areas are regions that balance their supply and demand independently; they can be connected through intertie lines, but still operate their own set of resources. Larger areas will see less variability compared to small balancing areas, making it advantageous to optimize over one large area [9]. The trend toward making larger balancing areas will allow more efficient utilization of resources (e.g., scheduling of reserves, unit commitment, etc.), decrease peak generation requirements, and increase the minimum load level. However, there are trade-offs that are made when network size increases. This chapter focuses on economic trade-offs, but enlargement also involves computational challenges. Larger system sizes create bigger models with more nodes and variables, which make incorporating additional complexities into system operations difficult. Furthermore, as system size increases, it becomes more difficult to solve scheduling models to optimality (larger duality gaps, etc. become necessary because of the model size), so the putative efficiency improvements of enlarging balancing areas might not be realized.

Consolidating balancing areas will increase the number of variable resources and the extent of the geographic area where the resources are located. While this might reduce the impact of output variability, it might not significantly impact the effects of forecast uncertainty. Renewable energy will produce uncertainty in the day-ahead forecast, which will increase as the penetration of renewables grows. In order to assess the impacts of balancing area consolidation in the presence of a high penetration of renewable energy, models for resource scheduling are created and compared. By comparing market models from different time frames, it can become apparent which types of coordination provide the greatest benefits and lowest costs. The questions of coordination and consolidation are especially relevant in the Western U.S. today; there are proposals for a regional operator [177], at the same time there are proposals for new balancing areas [192].

The second half of the chapter concentrates on balancing area coordination considering the complicating factor of externalities, particularly, cases when each balancing area values an externality asymmetrically. An externality is a cost or consequence that is incurred outside of the market setting. Environmental externalities can be challenging to price. If they are local like particulate matter pollution, the surrounding community is harmed. However, externalities like carbon emissions from power plants do not stay local and are not necessarily created locally; they impact the region and even global communities. If one balancing area aims to reduce its consumption of carbon emissions, it must consider all imports, and not solely the power plants in the target region. The challenge becomes ensuring imports do not also increase emissions, especially if the exporting balancing area does not aim to reduce emissions to the same degree.

Similar to the first half of the chapter, the Western U.S. is an ideal case study. California has implemented a cap-and-trade system for carbon emissions while also completely consolidating the real-time market with surrounding balancing areas. These regions have not implemented a carbon reduction scheme, but still want to sell to California customers. Several proposals for coordinating the trade have been proposed, and each will be evaluated for its ability to reduce carbon emissions, maximize market surplus, and price energy consistent with economic incentives. The proposals are compared against a first-best and do-nothing model.

This chapter contributes to current literature through modeling of balancing area coordination in time – between market auctions, and in space – through incorporation of externalities. Simulations in the first half reveal that increased trade in one market does not mean overall economic efficiency increases; all temporal interactions must be evaluated. Market models in the second half judge proposals for incorporating carbon emissions, finding all have downsides without any one dominating for both cost and emissions reduction. Both simulations use hurdle rates to model trade, demonstrating their simplicity as a modeling tool and disadvantage as a means to capture complexity.

5.2 BACKGROUND AND LITERATURE REVIEW

Research on balancing areas, and more generally seams issues, is an ongoing and prevalent topic in the power systems literature [193]. Although some issues are specific to particular areas or regions, many issues span countries and markets. Much of the literature can be divided into those that accept a means of coordination and simulate policies or assess costs, and those that attempt to optimize coordination. This project chooses to accept the current coordination mechanisms as given, and simulate different schemes.[35] Literature on each topic shows the depth of complexity that exists at the seams.

With increased integration of renewable energy, many in the literature have turned to balancing area coordination. There are many types of coordination among regions, with two extremes being complete consolidation and no trade or harmonization. The former refers to two areas joining or merging together; in an electricity market context, it entails optimizing both networks together in a single optimization problem. As an example, in 2013, Entergy joined the Midcontinent ISO [194]. MISO was then tasked with the joint optimization of the Entergy utility regions along with their existing footprint.

While there are many economic rationales for consolidation, renewable energy is increasingly encouraging regions to consider the benefits of consolidation, or at least further coordination. An early review paper on wind integration methods referenced balancing area coordination and consolidation as being important to minimizing the cost of renewable integration [136]. They suggest that increasing the size of balancing areas will help small systems balance variability, and will lead to fewer hours when expensive peaking units are needed. The proposal did not specify how areas should consolidate; however [195] emphasizes that consolidation does not need to

[35] The later area of literature considers the different modeling techniques to improve passing information across seams. Seminal work in the area uses decomposition techniques proposed by Kim and Baldick [242], [243] for solving large scale optimal power flow problems (OPF). Additional work of interest can be found in [244] and [245], which focus on the type and amount of information that must be passed between regions in order to determine a co-optimized dispatch.

solely come from physical consolidation, but also from what they call virtual consolidation, meaning sub-hourly scheduling, preforming economic dispatch in larger regions, and dynamic scheduling. In discussing the key drivers for large-scale wind integration, [196] also argues for better inter-regional coordination, which includes ancillary service markets, imbalance settlements and capacity calculations. An NREL report on the Western Interconnection showed that when balancing areas are aggregated, both the average and the maximum ramping needed are significantly less than each area individually [88]. Kirby and Milligan have written a series of papers concerning balancing areas as it relates to wind integration: analyzing ramping capability in regions with high penetration [197], utilizing sub-hourly markets to ease integration [198], and analysis of results in a 2008 NREL report [199]. Although the conclusions are insightful, these studies differ in methodology from this project, as they do not perform a rigorous power flow analysis of the consequences.

The literature specific to the European network is extensive, largely due to the continual integration of the European day-ahead market. An examination of centralized market clearing for coordinating power exchanges and pools in Europe can be found in [200], [201] and [202]. The authors in [203] describe problems that arise when using an LMP through simple market coupling examples, such as price ranges and single prices in networks without congestion. An additional detailed discussion of the European market coupling and modeling, especially considering the cooperation of the transmission system operator, can be found in [204]. The impacts of renewable variability across Europe are analyzed in [205], where the impacts of increasing transmission infrastructure vary depending on the penetration of renewables. The research on integration in Europe is often specific to a set of countries; for example, studies on coordination studies have been done for Northern Europe [113] and the Netherlands/Belgium [206]. Although their study is based in Europe, the authors in [207] propose a generalized day-ahead and real-time coordination model to analyze the interactions between regions, solved as stochastic mixed complementarity problem.

While the comparisons are similar to those proposed in this project, the solution method and analysis differ. Both compare two-period two-zone models, and this project further extends that analysis to include impacts of renewable policies in each region and uses modeling that is not just theoretical, but could currently be implemented by a system operator. This project differs from general balancing literature because it uses a power flow market model to assess the impacts of coordination. It also contrasts from much of the European literature because it explores the impacts of renewable policies and specific U.S. inter-regional barriers. There is a distinct difference between European market design and markets run by regional transmission operations in the U.S.; rigorous analysis has not been previously performed considering characteristics of a U.S. market, such as nodal pricing and real-time coordination of markets.

5.3 BENEFITS OF SPATIAL AND TEMPORAL COORDINATION

Combining balancing areas will increase the number of variable resources and the geographic area where the resources are located. While the resulting diversity of sources and loads might reduce the impact of output variability, it does not necessarily reduce the economic costs of forecast uncertainty. Day-ahead forecasts of renewable energy are uncertain, and this uncertainty will increase as the penetration of renewables grows [208]. This section assesses costs and prices for different levels of coordination between adjacent balancing areas, and also considers the influence of variable renewables and flexible resources on system costs. In particular, the question addressed here is:

What are the economic and reliability-related benefits and costs of alternative approaches to coupling neighboring balancing areas, ranging from complete merger to partial coordination of electricity markets?

Three different levels of coordination are compared for both day-ahead and real-time markets, which are the two most common types of power markets. Most coupling proposals address one or the other those market types. Complete consolidation and minimal coordination are compared against the use of a hurdle rate to control trade between the regions. This comparison on a test (hypothetical) system will reveal the drawbacks, both economic and reliability-related, of each coordination type, and act as a basis for further comparisons of renewable policies. The two models used for the day-ahead and real-time market are described below.

The unit commitment (UC) model commits and schedules all generation resources for each hour in the following day and is sometimes used to model several weeks of commitments. Generator constraints are included in the UC model to allow for the binary on/off commitment decisions. Each generator submits a two-part or three-part bid, where one part is the marginal cost of operating the generator and the other two are fixed costs, such as the expense of starting up and maintaining minimum load. Variable resources submit estimated production output, and system operators often input additional forecasts to reduce uncertainty. After the UC is run, the system operator will announce the committed generators and the resulting prices. The mathematical formulation for the day-ahead model is a typical unit commitment model with generator characteristics and network constraints, the formulation is in Section 5.4.2.a.

The real-time (RT) or dispatch model (called a balancing model in the European context) dispatches generators for the following 5-minute period up to several hours. Depending on the type of resources available, the model can commit fast-start units and dispatch resources such as demand response and energy storage. Units that have been previously committed can be dispatched up or down, and their cost functions now only reflect the marginal cost of delivering power without the fixed cost components. Variable resources can submit updated forecasts, as the market is closer to real-time; although day-ahead forecast errors will be higher, real-time forecasting errors exist as well. Like the day-ahead market, the simulation also outputs a commitment schedule, dispatch level, and set of prices for each node. The mathematical formulation for the real-time market makes modifications to the day-ahead market, and similarly does not contribute to literature or understanding of coordination research; the model can be found in Section 5.4.2.b. The following section describes additions made to the two-market models in order to simulate coordination.

5.4 DAY-AHEAD AND REAL-TIME MARKETS

5.4.1 Description

In order to analyze the economic impact of coordination between markets, three different simulations are proposed for each time scale (day-ahead commitment and real-time dispatch/short-start commitment) involving varying degrees of cooperation: complete coordination (consolidation), trade with hurdle rates, and minimal coordination (described in detail below). Each of these three types can be implemented in either the day-ahead or real-time markets, giving combinations of simulations, as seen in Table 5-1, which also gives actual examples of some of the combinations. The simulations involve both the day-ahead and real-time markets, simulating how day-ahead decisions constrain real-time decisions, creating nine models in total, one for each possible combination.

For U.S. markets, the simulation diverges from actual operation in several ways. First, there is a simple reserve requirement constraint rather than a side ancillary services market with multiple products (spinning reserves, non-spinning reserves, regulation up and down, and/or flexible ramping product). Second, there are no virtual bidders in either market, which can provide day-ahead/real-time arbitrage; as a result, expected day-ahead can diverge from real-time prices, which can lower market efficiency [209], [210]. Therefore, it is possible that some of the inefficiencies that are found with certain combinations might be at least partially mitigated by virtual bidding.

Table 5-1 Coordination between select markets in the U.S. and Europe

		Real-Time		
		Minimal coordination	Hurdle Rate	Full Integration
Day-Ahead	Minimal coordination	ERCOT & Rest of U.S.		
	Hurdle Rate	Most of Western U.S., EU non-market splitting	Regional authorities	CAISO EIM
	Full Integration	EU market splitting/coupling (EUPHEMIA)	Nordpool	Consolidation, Single RTO

These degrees of coordination mimic existing markets, which are contained in the boxes of Table 5-1. For instance, the California Independent System Operator has a fully integrated real-time market with many of its neighbors, called the Energy Imbalance Market (EIM). However, the day-ahead markets of these balancing authorities are not presently integrated and rely on rules and bilateral contracts to coordinate trade; therefore, it falls into the hurdle rate day-ahead / full integration real-time box. In Europe, markets are coupled in day-ahead, with a central market mechanism called EUPHEMIA [211]. In real-time, each transmission system operator balances their own region independently, which puts them in the full-integration day-ahead / minimal coordination real-time box. Some of the models are not necessarily practiced; for instance, if two markets have minimal coordination in the day-ahead market, it might be impractical to then fully integrate in real-time. This list in Table 5-1 not exhaustive but provides some real-world relevance to the simulations since most markets trade to some degree with their neighbors and/or use bilateral contracts, at least day-ahead.

The models can be simulated for any number of regions or countries. Many regions, especially in Europe, must balance interactions with many neighbors. In the U.S., there are often pairs of regions interacting due to the size of markets. New England-ISO shares a U.S. border with only New York ISO; California shares many state borders, but they are the only organized auction market (at present) in the Western U.S. This chapter chooses to model only two regions so that the outcomes and benefits can be easily distinguished. The following subsection describes the mathematical formulations for a basic unit commitment and real-time market model.

5.4.2 Mathematical Formulations

5.4.2.a Basic Day-Ahead Unit Commitment Market Model

Equations (5-1)–(5-22) are considered the basic day-ahead unit commitment model for this project and will be used as the base model for the various comparisons. The basic formulation is based on [212], [213], with minimum up time constraints and ramping constraints from [214]. The objective of the model is to minimize operating costs, which consist of fixed costs for startup and no-load conditions, as well as the marginal cost for the power dispatched. In this basic UC model, a simple objective is used; however, more complex formulations can be used depending on the cost function of the generators. Constraints (5-2)–(5-4) limit the power capacity, line generation limits, and voltage angle limits respectively. The dc power balance equation is shown in (5-5), which is a linear approximation of the ac real power flow equation. The origin and destination nodes are represented by *m* and *n* respectively.[36] Constraints (5-6) and (5-7) define the startup and commitment status of the generator, where the variable *v* can either be 1 (startup occurs in that interval) or 0 (otherwise). The ramping capability of the generators is defined in (5-8) and (5-9), where (5-8) limits ramping up and (5-9) limits ramping down. Constraints (5-10) and (5-11) define the minimum up and down times for the generators during the 24 hour commitment period.

---

[36] The real power transferred on line *k* from bus *m* to bus *n* is $P_k = V_m^2 G_k + V_m^2 G_m - V_n V_m (G_k \cos(\theta_m - \theta_n) + B_k \sin(\theta_m - \theta_n))$. For the linearized expression (dc power flow), we assume voltage magnitudes are close to one per unit, the shunt conductance is negligible, the susceptance is defined as $B_k = -1/X_k$, and the voltage angle difference on the line between the buses is small so that $\sin(\theta_m - \theta_n) \approx \theta_m - \theta_n$ and $\cos(\theta_m - \theta_n) \approx 1$. The real power equation can then be reduced to $P_k = G_k - G_k - B_k(\theta_m - \theta_n) = B_k(\theta_n - \theta_m)$. Many textbooks provide some explanation of the dc power flow, I find the book appendix chapter by Seifi and Sepasian to be the most similar to the notation I prefer [246].

Similar to the dispatch model, the wind injection can be curtailed in the node balance constraint (5-12), and is limited by the day-ahead forecast in (5-13). Demand can also be shed at a high cost, shown in (5-14), often called the value of lost load (VOLL). The reserve capacity of each generator in the system is shown in (5-15) and (5-16) for spinning reserve, and (5-17) for non-spinning reserve. Constraints (5-18)–(5-19) define the reserve requirements for spinning and non-spinning reserves. The first requirement is based on the NREL 3+5 rule, where 3% of demand and 5% of variable generation is required. The two percentages are represented by α and β, which can be adjusted as necessary depending on the system size, penetration, and risk aversion. The second requirement states there must be enough reserve to withstand the loss of the single largest generator A small percentage of wind is also included in these constraints to ensure enough reserve is available in case of a wind contingency. Spinning requirements must account for half of each rule (5-18) and (5-20), while spinning and non-spinning together account for the full reserve requirement (5-19) and (5-21). Finally, the commitment variable defining the status of a generator is restricted to be a binary variable in (5-22). The variable and parameter definitions can be found in the Notation section at the beginning of the dissertation.

$$\min \sum_{\forall t} \sum_{\forall g} \left(c_g p_{g,t} + c_g^{\mathrm{SU}} v_{g,t} + c_g^{\mathrm{NL}} u_{g,t} \right) + \sum_{\forall i} VOLL\left(d_{i,t}\right) \tag{5-1}$$

Subject to

$$p_{g,t} \geq u_{g,t} P_g^{\min}, \ \forall\, g, t \tag{5-2}$$

$$F_k^{\min} \leq f_{k,t} \leq F_k^{\max}, \ \forall\, k, t \tag{5-3}$$

$$-\theta^{\max} \leq \theta_{i(k),t} - \theta_{j(k),t} \leq \theta^{\max}, \ \forall\, k, t \tag{5-4}$$

$$f_{k,t} = B_k\left(\theta_{n(k),t} - \theta_{m(k),t}\right), \ \forall\, k, t \tag{5-5}$$

$$v_{g,t} \geq u_{g,t} - u_{g,t-1}, \ \forall\, g, t \tag{5-6}$$

$$0 \leq v_{g,t} \leq 1, \ \forall\, g, t \tag{5-7}$$

$$p_{g,t} - p_{g,t-1} \leq R_g u_{g,t-1} + R_g^{\mathrm{SU}}\left(v_{g,t}\right), \ \forall\, g, t > 1 \tag{5-8}$$

$$p_{g,t-1} - p_{g,t} \leq R_g u_{g,t} + R_g^{\mathrm{SD}}\left(1 - u_{g,t}\right), \ \forall\, g, t > 1 \tag{5-9}$$

$$\textstyle\sum_{r=t-\tau_g^{\mathrm{UT}}+1}^{t} v_{g,r} \leq u_{g,t}, \ \forall\, g, t \geq \tau_g^{\mathrm{UT}} \tag{5-10}$$

$$\textstyle\sum_{r=t+1}^{t+\tau_g^{\mathrm{DT}}} v_{g,r} \leq 1 - u_{g,t}, \ \forall\, g, t \leq |T| - \tau_g^{\mathrm{DT}} \tag{5-11}$$

$$\textstyle\sum_{\forall k \in \delta^+(i)} f_{k,t} - \sum_{\forall k \in \delta^-(i)} f_{k,t} + \sum_{\forall g \in g(i)} p_{g,t} + w_{i,t}^{\mathrm{inj}} = d_{i,t}, \ \ \forall i, t \tag{5-12}$$

$$0 \leq w_{i,t}^{\text{inj}} \leq W_{i,t}, \; \forall\, i,t \tag{5-13}$$

$$0 \leq d_{i,t} \leq D_{i,t}^{\max}, \; \forall\, i,t \tag{5-14}$$

$$r_{g,t}^{\text{SP}} \leq u_{g,t} P_g^{\max} - p_{g,t}, \; \forall g,t \tag{5-15}$$

$$r_{g,t}^{\text{SP}} \leq R_g^{10} u_{g,t}, \; \forall g,t \tag{5-16}$$

$$r_{g,t}^{\text{NSP}} \leq R_g^{10}\left(1 - u_{g,t}\right), \; \forall g,t \tag{5-17}$$

$$\sum\nolimits_{\forall g} r_{g,t}^{\text{SP}} \geq \tfrac{1}{2}\left[\alpha \sum\nolimits_{\forall i} d_{i,t} + \beta \sum\nolimits_{\forall i} W_{i,t}\right], \; \forall t \tag{5-18}$$

$$\sum\nolimits_{\forall g}\left(r_{g,t}^{\text{SP}} + r_{g,t}^{\text{NSP}}\right) \geq \alpha \sum\nolimits_{\forall i} d_{i,t} + \beta \sum\nolimits_{\forall i} W_{i,t}, \; \forall t \tag{5-19}$$

$$\sum\nolimits_{\forall g'} r_{g',t}^{\text{SP}} \geq \tfrac{1}{2}\left[p_{g,t} + r_{g,t}^{\text{SP}} + \beta \sum\nolimits_{\forall i} W_{i,t}\right], \; \forall g,t \tag{5-20}$$

$$\sum\nolimits_{\forall g'}\left(r_{g',t}^{\text{SP}} + r_{g',t}^{\text{NSP}}\right) \geq p_{g,t} + r_{g,t}^{\text{SP}} + \beta \sum\nolimits_{\forall i} W_{i,t}, \; \forall g,t \tag{5-21}$$

$$u_{g,t} \in \{0,1\} \tag{5-22}$$

5.4.2.b Basic Real-Time Market Model

Equations (3.0)-(3.11) are a basic economic dispatch model or an imbalance model. This program is used for real-time simulations and fixes the commitments of the slow generators from the day-ahead model ($u_{g,t}^{*}$). The fast generators are not fixed, and allowed to turn on during the day, for example, 15 minutes ahead of being dispatched. The objective of the model minimizes dispatch costs, which consist of the product of the linear cost and the power dispatched. Constraints (3.1)-(3.3) limit the power capacity, line generation limits, and voltage angle limits respectively. The dc power balance equation is shown in (3.4), which is a linear approximation of the ac real power flow equation. Constraints (3.1)-(3.4) are the same as (1.1)-(1.4) in the UC model. The ramping capability of the generators is defined in (3.5). Since the commitment of the slow generators is fixed, (3.6) specifies that the commitment variable cannot vary from the fixed schedule. The wind generation is deterministic, since the model mimics real-time; however, wind curtailment is allowed. The node balance constraint is in (3.8). Constraints (3.9) and (3.10) limit the wind injected into the system, $w_{i,t}^{\text{inj}}$, and the demand between zero and their respective maximums, forecast and load. Finally, the commitment status variable is constrained between 1 and 0 in (3.11); this will only affect the fast generators.

$$\min \sum_{\forall t} \sum_{\forall g} c_g p_{g,t} + \sum_{\forall g \in G^{\text{fast}}} c_g^{\text{NL}} u_{g,t} + \sum_{\forall i} VOLL(d_{i,t}) \tag{5-23}$$

Subject to

$$u_{g,t} P_g^{\min} \le p_{g,t} \le u_{g,t} P_g^{\max}, \ \forall\, g, t \tag{5-24}$$

$$F_k^{\min} \le f_{k,t} \le F_k^{\max}, \ \forall\, k, t \tag{5-25}$$

$$\theta^{\min} \le \theta_{i(k),t} - \theta_{j(k),t} \le \theta^{\max}, \ \forall\, k, t \tag{5-26}$$

$$f_{k,t} = B_k\left(\theta_{n(k),t} - \theta_{m(k),t}\right), \ \forall\, k, t \tag{5-27}$$

$$-u_{g,t} R_g \le p_{g,t} - p_{g,t-1} \le u_{g,t} R_g, \ \forall\, g, t \tag{5-28}$$

$$u_{g,t} = u_{g,t}^*, \qquad \forall g \in G^{slow}, t \tag{5-29}$$

$$\textstyle\sum_{\forall k \in \delta^+(i)} f_{k,t} - \sum_{\forall k \in \delta^-(i)} f_{k,t} + \sum_{\forall g \in g(i)} p_{g,t} + w_{i,t}^{\text{inj}} = d_{i,t}, \ \ \forall i, t \tag{5-30}$$

$$0 \le d_{i,t} \le D_{i,t}^{\max}, \ \forall\, i, t \tag{5-31}$$

$$0 \le w_{i,t}^{\text{inj}} \le W_{i,t}, \ \forall\, i, t \tag{5-32}$$

$$u_{g,t} \in \{0,1\}\ \forall g, t \tag{5-33}$$

5.5 SIMULATIONS OF COORDINATED MARKETS

5.5.1 Simulation 1: Consolidation

As an ideal benchmark for efficiency, the first simulation treats the two systems as one single balancing area, which would apply to both types of markets. This would occur if two adjacent regions combined balancing needs for all generation and load, and decided to schedule resources together in either the day-ahead or real-time. The Energy Imbalance Market, which includes the California ISO and several neighboring utilities, is such an example for the real-time market. The model would consist of the objective and constraints in the Appendix (0, equation numbers below refer to the equations in that appendix). The generator set would include generation from both areas as a single input and all intertie lines would allow power to flow following Kirchhoff's laws and the transmission line limit. The output of the model would show the schedule and operating cost for a single deterministic balancing area.

The below outlines in schematic form the two versions of this model

Consolidation Model

Day-ahead version	*Real-time version*
Objective (5-1)	Objective (5-23)
Constraints (5-2)-(5-22)	Constraints (5-24)-(5-33)

5.5.2 Simulation 2: Minimal Coordination

The second simulation models a case where the two areas are solved with minimal coordination. Although the worst case scenario would likely result from solving the two areas as completely independent (i.e., open breakers on the intertie lines), that is not realistic for the present power system. The minimal coordination case likely the best case without direct trade, since there is some cooperation between where flows occur. The network modeling assumed ac linkages, therefore coordination within one of the three interconnections in the U.S.[37] Although dc linkages could be modeled, the use of dc lines can be directed, and would require different assumptions about contracts within the region.

The model is similar to the basic UC with the addition of a constraint and a change to the reserve. The additional constraint, (5-34), limits the sum of the flows on the intertie lines to be zero; this same constraint is added to the real-time model (notation can be found at the beginning of the dissertation in a section called Notation). Power can flow between the two

[37] The U.S. is divided into three major interconnections, meaning all power operates at the same frequency within the region [247]. The Western Interconnection reaches from Canada to Mexico, from the Pacific Ocean to the Rocky Mountains. The Eastern Interconnection is comprised of many sub-regions for reliability, also extending into Canada and from the Atlantic Ocean to the Rockies. Finally much of Texas is electrically isolated, with a single market and balancing authority called ERCOT.

systems on individual lines, but the net across all lines for each hour must sum to zero. Any flows due to previous transactions would also necessarily sum to zero, leaving a zero net imbalance. The reserve requirements for this system are determined for each area individually. Instead of a system wide requirement, the amount of additional contingency reserve is determined with requirement by area. The requirement itself does not change, but the generators available to meet the requirement are limited by area, as was examined in Section 3.3.2.d.

$$\sum_{\forall k \in K^{\mathrm{IT}}} f_{k,t} = 0, \;\; \forall t \tag{5-34}$$

Minimal Coordination Model

Day-ahead version	*Real-time version*
Objective (5-1)	Objective (5-23)
Constraints (5-2)-(5-22),(5-34)	Constraints (5-24)-(5-33),(5-34)

5.5.3 Simulation 3: Trade with Hurdle Rates

5.5.3.a Factors that Lead to the Use of Hurdle Rates

In reality, there are many types and degrees of coordination beyond consolidation and the minimal coordination model. Inter-region coordination can occur between individual parties or through the system operator (a market operator or a utility). Individual participants in a balancing area can create bilateral contracts with one another to hedge the risk of day-ahead and real-time price spikes. These contracts are private information, and usually extend for long terms. For example, a solar farm in Nevada can sell its output to a city in California for the first ten years after it is commissioned; agreements like this are known as Power Purchase Agreements, and can help finance renewable projects. On the demand side, a utility that wants to ensure customers have a stable price might sign a bilateral contract with a supplier who then bears the risk of price volatility and provides power through a mix of its own output and purchases from the spot and bilateral bulk markets. These contracts tend to make up a large part of the energy market; in 2016, the PJM market was composed of 12.9% bilateral contracts [215].

From the system operator perspective, there are several opportunities for coordination with neighboring regions. As a means of partial consolidation, operators can establish transfers through dynamic scheduling or pseudo-tie lines [216]. When a resource is physically located outside of a balancing area, it can serve load in that balancing area by using a dynamic schedule. The dynamic schedule depends on communication and metering between the areas to account for changes in energy. If no physical ties exist between the resource and the balancing area, a pseudo-tie line can be used to transport energy. It relies on adjusting controls between the two areas, called Area Control Error (ACE), called Area Control Error (ACE), to account for changing energy use. The resource is used as if it was part of the target balancing area; in this way, a wind farm in Montana to be dispatched to serve California load as if it were a California resource.

Both bilateral contracts and operator actions can cause issues because they rely on availability of inter-tie capacity. Each has a means of reserving and paying for transmission across the network, but the use of these different products can cause complications when modeling. To ease the strain of modeling many types of products and reservations on lines, hurdle rates can be used as a proxy.

5.5.3.b Defining Hurdle Rates

Modeling trade between neighbors can be difficult for both technical and political reasons. The model must make assumptions about how much information is shared between the two areas and the depth of detail. Realistic modeling might take many iterations to determine an appropriate hurdle rate, which can be time intensive for an operational market. Each BA can also have its own set of rules, making a generic model nearly impossible. Due to the many complexities that arise between regions, most models choose to model inter-regional trade with a hurdle rate [217]. A hurdle rate is a simple price per megawatt that is placed on the intertie line(s) from one region to another. The price is determined exogenously, which can act as "resistance" to flow on the line; it represents the minimum price difference that will cause power to flow on the line or a minimum benefit for a transaction to take place.

A hurdle rate can be simply modeled with an addition to the objective and constraint (5-37), where S_t^{AB} and S_t^{BA} are the net flow on the intertie lines in one direction, and H_t is the hurdle rate. The initial simulations used a $3/MWh hurdle rate, which is in the range of hurdle rates used in power flow simulations of the West [170], [218] (which can range from $0.53/MWh to $14/MWh). The same constraint flow constraint in (5-37) is added to the real-time model.

$$\min \sum_{\forall t} \sum_{\forall g} (c_g P_{g,t} + c_g^{\mathrm{SU}} v_{g,t} + c_g^{\mathrm{NL}} u_{g,t}) + \sum_{\forall i} VOLL(d_{i,t}) + \sum_{k \in IT} H_t(s_t^{\mathrm{AB}} + s_t^{\mathrm{BA}}) \tag{5-35}$$

$$\min \sum_{\forall t} \sum_{\forall g} c_g p_{g,t} + \sum_{\forall g \in G^{fast}} c_g^{\mathrm{NL}} u_{g,t} + \sum_{\forall i} VOLL(d_{i,t}) + \sum_{k \in IT} H_t(s_t^{\mathrm{AB}} + s_t^{\mathrm{BA}}) \tag{5-36}$$

$$s_t^{AB} - s_t^{BA} = \sum_{\forall k \in K^{\mathrm{IT}}} f_{k,t}, \forall t \tag{5-37}$$

Hurdle Rate Model

day-ahead objective (5-35)	real-time objective (5-36)
day-ahead constraints (5-2)-(5-22), (5-37)	real-time constraints (5-24)-(5-33), (5-37)

5.6 RESULTS

A basic unit commitment model was created using AIMMS modeling software, version 3.13. The model is a deterministic model with two balancing areas, and a total of 73 buses. The input electrical system for the network is adapted from the three-zone Reliability Test System 1996 from [219], shown in Figure 5-1. The wind input data is developed from historical wind speed data from NREL wind site 3776 in California [220]. The bus, generator, and load data are in the

Appendix. A total of nine simulations were run, including all combinations described in Table 5-1.

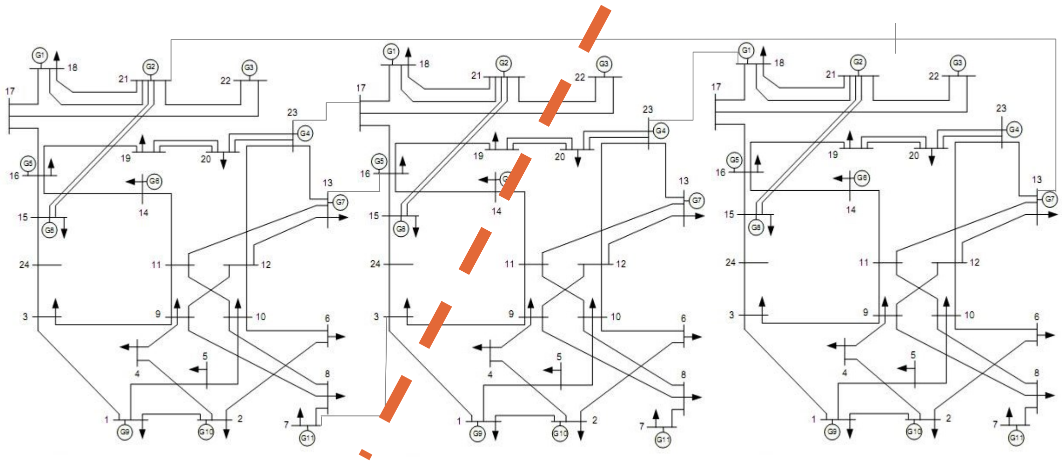

Figure 5-1 IEEE RTS 96 test case

The total cost of the IEEE Test System can be found in Table 5-2. The hurdle rate and full integration models (lower 2x2 quadrant) were almost the same, resulting in costs within the MIP gap. The cases where each market was not coordinated in the day-ahead but was coordinated in real-time (first row) resulted in higher costs, the highest being when real-time was also not coordinated. The worst cases of the nine resulted when there was minimal coordination in real-time, but some coordination in the day-ahead market (either a hurdle rate or consolidation). These cases resulted in a large amount of load shedding, leading to a considerable amount of lost load costs in the objective. Note that these cases would have been infeasible without including load shedding as a variable.

A similar result occurs when we look at average prices over the 24-hour time period from the day-ahead and real-time markets in Table 5-3. The hurdle rate and integration cases, in any combination, result in the same averages over the course of the day for each of the markets. Similarly, when there is minimal coordination in the day-ahead market and any amount in real-time, the average prices remain the same for each of the markets. The significant difference occurs when there is minimal coordination in real time and any amount of coordination in the day-ahead market. Again, the models would have been infeasible without lost load. With a high cost of lost load (often referred to as the value of lost load or VOLL), the prices reflect the demand that cannot be met.

Models using a higher hurdle rate are simulated, and the resulting prices for day-ahead and real-time are in Figure 5-2. The diagonal simulations are compared (minimal coordination in both day-ahead and real-time, full integration in both, etc.) against hurdle rates of \$6/MWh and \$10/MWh. Both increased rates had little impact on prices, lowering the percent change between day-ahead and real-time by a percentage.

Table 5-2 Real-time total costs over 24 hours (million$)

		Real-Time		
		Minimal Coordination	Hurdle Rate	Full Integration
Day-Ahead	Minimal Coordination	3.07	2.73	2.73
	Hurdle Rate	292	2.18	2.18
	Full Integration	308	2.18	2.17

Table 5-3 Day-ahead and real-time prices (\$/MWh)

		Real-Time		
	DA/RT	Minimal coordination	Hurdle Rate	Full Integration
Day-Ahead	Minimal coordination	52 / 47	52 / 47	52 / 47
	Hurdle Rate	45 / 5240	45 / 32	45 / 32
	Full Integration	45 / 6214	45 / 32	45 / 32

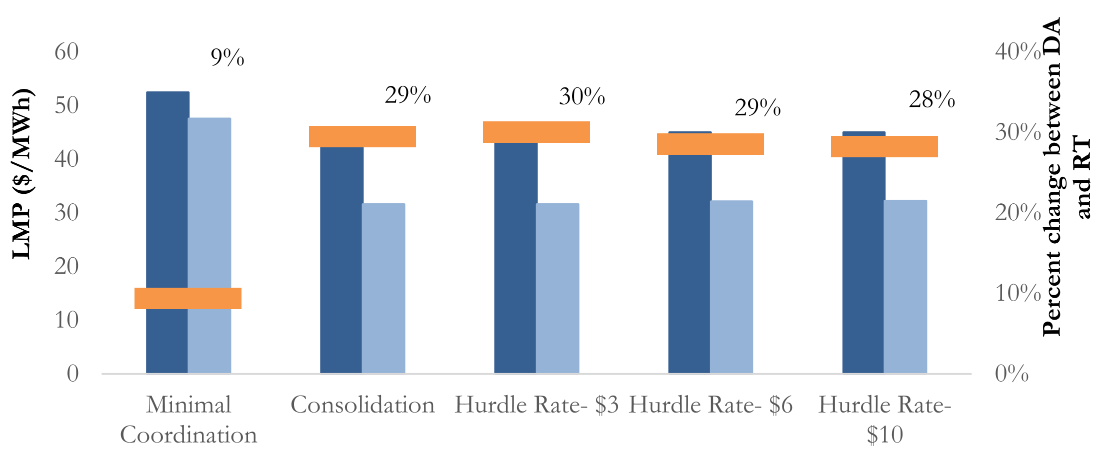

Figure 5-2 Change between day-ahead and real-time prices

Finally, total costs can be compared against the total amount of trade occurring throughout the two regions. Figure 5-3 shows the nine simulations with total cost in $ and trade in MWh values for each, excluding the cost of load shedding (assuming emergency generators where turned on in response to added need). The day-ahead cost/trade value is shown as a square for each of the three types of coordination. The triangles show full integration, and have the lowest cost and highest trade values in the comparison. The hurdle rate simulation is close behind or takes on the same value in the case of minimal coordination in day-ahead. The figure shows even with some kind of coordination in real-time, the lack of coordination in day-ahead leads of higher costs and less trade.

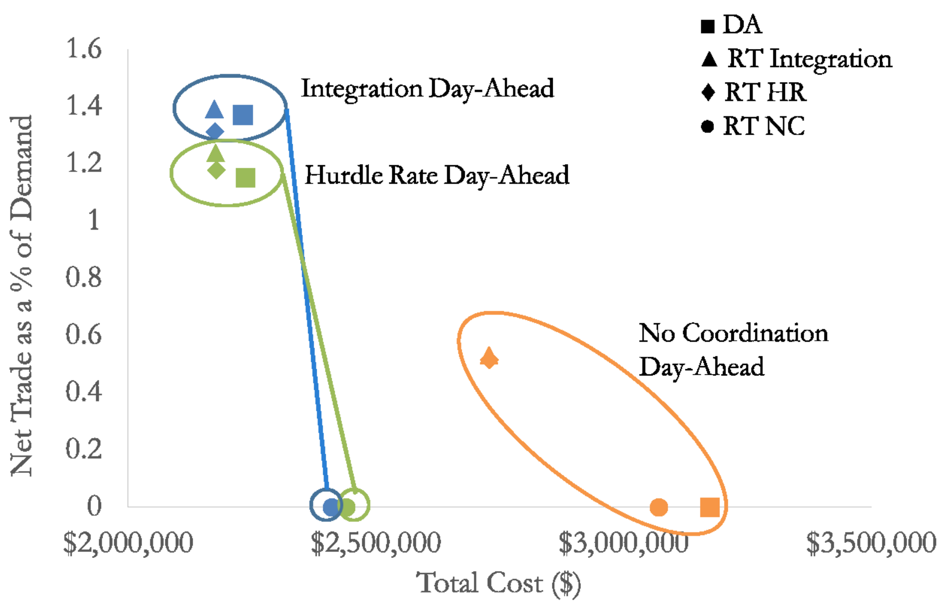

Figure 5-3 Total costs compared net trade as a percent of demand

5.7 DISCUSSION

This chapter analyzes trade and coordination between neighboring balancing areas. The first part, Section 5.3, examines the interactions between balancing areas in time (through the day-ahead and real-time markets) and space (through a spectrum of coordination). The second part, Section 4.2.1, evaluates models for trading externalities (emissions) in a spot market. Each compares different models for trade; the first part aims to find inefficiencies between markets, while the second focuses on a second-best model for trade considering externalities. The simulations use simple models for ease of understand the results, and the lessons might show broad trends, but are also dependent on the system and data utilized.

Trade between balancing areas can cause issues for most regions throughout the world with liberalized markets. Frequent problems have arisen between countries in Europe and states in the U.S. because each region can establish different laws and policies. Those policies can vary depending on the time frame. As explained in Section 5.3, the rules for day-ahead markets can differ from the rules for real-time or balancing markets. In Europe, there is centralized market clearing in day-ahead and independent balancing in real-time. In the West the opposite occurs; the day-ahead market sees trade between regions and there exists a real-time fully consolidated market, the Energy Imbalance Market, operated by California ISO. While each region has progressed over time, each chose to coordinate inversely across time. Coordination will become more important as the penetration of wind energy grows. Spatial diversity is an important aspect for successful wind integration, and increased trade between regions can help accommodate wind.

Section 5.3 analyzes methods for coordination in time and with a simple model, finds that not all means of coordination produce the same results. Results from the IEEE Test Case find the amount of coordination in the day-ahead market has a direct impact on the efficiency of the real-time market. With either full integration or use of a hurdle rate in the day-ahead market and either type of coordination in real-time, the market outcomes are very similar. Both sets of simulations see similar costs and prices.

However, with some kind of coordination in day-ahead and minimal coordination in real-time (like some areas in European), the real-time simulation for the Test Case is infeasible. As the model allows for load shedding, costs skyrocket due to the high value placed on shedding customers. This simple case shows manual commitment of generators or used of expensive emergency generation would be necessary since there is not enough generation online. Since there are not massive blackouts in regions with this trading scheme, these results do not necessarily scale to real systems. They are a reminder for policy makers that decisions to coordinate trade in one time scale might benefit that market, but might not benefit the system as a whole. It is important to evaluate the impacts markets can have on each other, rather than a single market in isolation.

CHAPTER 6
CONCLUSIONS

The four projects that make up this dissertation attempt to identify and improve inefficiencies in the electric power grid. Some of the inefficiencies are due to existing market issues while others have arisen due to the influx of renewable energy and the implementation of policies addressing greenhouse gas emissions. The electric power system is extremely complex and making any one change can have great and unexpected impacts on other parts of the grid. Chapters 2, 3, 4, and 5 each take a complex market or operational issue and use optimization-based operations planning and market simulation models to suggest solutions and efficiency improvements.

Chapter 2 begins with a basic model of day-ahead system operations and explains how even a simple model produces great complexity. Pricing from the unit commitment problem is not straightforward, and any change to current practice must consider how incentives for both supply and demand are impacted. The Dual Pricing Algorithm is a proposed model that attempts to satisfy certain economic criteria concerning efficiency and allocation of prices, while maintaining the efficient generator and demand schedules. Its development recognizes the need for rules to accompany any proposed pricing mechanism, since pricing rules should be enacted in cooperation with procedures ensuring incentive compatibility. The necessity for rules also impacts Chapter 3, where reserve coordination must follow a set of rules. The simulations in

Chapter 3 identified reserve coordination between countries as having the greatest net improvement compared to the other proposed adjustments. However, in order to appreciate the benefits of coordinated reserve procurement, each country must agree to a set of rules. Disagreement over rules or procedures can decrease trade, lessening the potential surplus due to coordination.

Chapters 3, 4, and 5 all compare proposed and existing models to identify their relative economic and efficiency benefits. The purpose of each chapter is different, but the evaluation metrics they share are financial benefits and integration of renewable energy. Chapter 3 finds coordinating reserves in the balancing or real-time markets of neighboring systems lowers system costs for the case study. Chapter 4 shows how modeling dispatch and environmental costs under alternative carbon and electricity trade policies can tradeoff emissions reductions for higher costs. Finally, the test case in Chapter 5 shows coordination in day-ahead markets without coordination in real-time might lead to higher system costs due to balancing actual operations in real-time.

Power system operations and markets tie each chapter together, where the unit commitment model is a linchpin for analysis for three out of the four chapters. Chapter 2 proposes a new method for pricing, using the simplest unit commitment problem in order to better analyze the impacts of modeling changes on prices. The unit commitment problem is what drives the new pricing proposal, which modifies the dual formulation of the classic unit commitment model in order to satisfy certain economic criteria. Chapter 4 also uses a simple model (dispatch only for a real-time market) to best evaluate how simple changes to pricing carbon between neighbors can impact costs and emissions. Chapters 3 and 5 use a more complex, dynamic unit commitment model including many generator characteristics to address system wide impacts of reserve and trade models respectively. Using an operations model in each chapter shows the importance of tailoring the model for the application at hand. Simple models can offer deep insights, and complex models allow for numerical answers in the context of practical applications.

Future work on the topics addressed in these chapters must account for technology innovations or market changes that will take place in the coming years and decades. Many tangible technologies will influence operation of the power grid; decreasing battery costs, smart grid devices to control ac power flow, and use of transmission switching might show significant improvements to power operations. Technology innovation can also greatly impact demand-side participation in markets. It is in its infancy, and extensive investments in physical technology and control software are underway to support further development in demand-side participation. Research into the role of aggregators has increased. Third-party companies might soon be able to submit a single offer on behalf of many smaller customers to wholesale market auctions, where participation from small customers might not be possible otherwise.

The concept of a distribution system operator (DSO), to complement the ISO, has also gained popularity [221]. A DSO operator can have greater insight into the state of the distribution system to either take a passive or active role in coordinating distributed energy resources, such as rooftop solar. Finally, the increased use of blockchain technology might increase the ease of bilateral trades between neighbors, allowing a secondary market for "local" renewable energy [248]. These ideas and technology can provide means for demand to understand and participate in wholesale markets, finally achieving the demand side participation conceived by Schweppe *et al.* in *Spot Pricing of Electricity* [13].

Thinking broadly about all market design issues raised in this dissertation, future work could reimagine existing market design; if a group of electricity experts started with a blank slate, would they imagine the market we have today? Without any existing infrastructure, what could electricity markets look like? Would this group still propose a wholesale spot and capacity market? Should a market with mostly zero cost resources look meaningfully different than the market today? These questions are of interest to many in the industry and would be a very worthwhile continuation of this research.

In the remainder of this chapter, I suggest some potentially useful and intellectually challenging research directions for each of the four problems addressed in my dissertation. Most suggestions are specific to the chapter's topic, but some can apply to electricity market research generally. For instance, Chapters 2, 4, and 5 would benefit from larger test cases. Simple examples are necessary to learning the fundamental attributes of a model, but larger more complex systems can elucidate the trends that can be seen over time or alternate participant behavior. The proposed future work for Chapters 2-5 are described in Sections 6.1-6.4 respectively.

6.1 FUTURE WORK ON PRICING

Electricity market pricing is complex and can often involve conflicting goals. Providing a transparent price might also incentivize generators to overbid. Ensuring non-confiscation of demand might interfere with incentive compatibility in elastic consumers, causing pay-as-bid incentives. The incentives for supply and demand given the complex cost structure of electricity and its lack of largescale economic storage make developing and implementing a single pricing method difficult. The method developed in Chapter 2, the Dual Pricing Algorithm, improves on several issues with current methods but does not solve all problems inherent in electricity pricing.

Since the design of electricity pricing is a multi-objective problem, it can be useful to elicit expert and stakeholder opinions regarding the relative importance of different aims (e.g., signals for investment, marginal cost pricing), and then apply those value judgments to the ranking of different proposals. This approach is not exclusive to Chapter 2, seeking expert feedback to rank methods can be applied to any of the market design questions addressed in this dissertation. As future work, such a multi-objective comparison of pricing system alternatives could be carried out with different ISO employees, generator operators, financial investment firms, demand aggregators, economists, and regulators. By explicitly considering and valuing tradeoffs, better justified recommendations concerning pricing mechanisms might emerge that are more consistent with stakeholder values [222], [223].

More specifically for the proposed DPA method, several studies can be carried out to continue the work. One question that has arisen concerns strategic bidding. An ideal pricing system would incent truth-revealing, i.e., truthful bidding of costs would be the dominant strategy. However, this is an unattainable ideal for practical pricing systems. New pricing methods will change the bidding behavior of the players, so the exploration of exactly how each pricing method changes bidding behavior would be of great value to the industry. This can be accomplished through use of a game-theoretic model in which bidding is a strategic variable. This might be structured as a mathematical program with equilibrium constraints [224], or by simulating the profit consequences of specific options for strategic bidding.

For long-term planning, future work on pricing can include a study of investment decisions given different pricing mechanisms [47]. In Chapter 2, one motivation for the DPA is increasing price transparency in order to send a stronger investment signal. This is based on the assumption that having more information about a pricing point of entry (or the lowest cost point at which a generator can enter a market and recover costs) will be a stronger signal than the marginal cost alone. As future work, it would be beneficial to simulate investment decisions under DPA, ELMP, and other proposed methods. As mentioned in the literature review, there have been some studies of the investment incentives for other pricing methods and replicating or advancing on one of these studies would help establish the long-term incentives of different proposals. Some new resources have high marginal costs and low or no fixed operating costs. For instance, the investment cost of batteries is high, and they are not yet considered economically viable at a large scale [225]. If prices were higher in certain locations, it might incent investment in such resources, especially if they can act as both a generation and consumption resource. Simulations using different pricing mechanisms might change welfare dynamics, and it will be important to assess who wins and who loses in the long-run.

Over the past several years I have thought a great deal about pricing and its interaction with economics and engineering. As engineers, we must learn the principles that economists hold dear, fundamental theories that govern the markets. The economists I speak with about pricing either shudder at the notion of fixed 'operating' costs, exclaim at the idea of moving away from marginal cost pricing, or throw up their hands and declare there might not be a supporting price for electricity. Many of the issues discussed in Chapter 2 can have theoretical solutions, but as an engineer, I am constantly concerned with implementation.

Given the constraints of the system as it stands in 2018, I am left with the belief that to uphold the principles in Chapter 2, pricing mechanisms need both an algorithm and associated rules. Rules, such as rules for uninstructed deviations described in Section 2.5.6, are needed to balance incentive incompatibility. For instance, a pricing method that moves away from marginal cost pricing might incentivize a generator to deviate from the optimal dispatch; when the price on its own does not incentivize staying on dispatch, a rule must be implemented to ensure compatibility. Any move away from marginal cost pricing raises concerns (and often panic). Because of the non-convexities in electricity markets, a pricing algorithm without rules is not likely to be enough to ensure markets remain incentive compatible for both supply and demand. As new proposals arise, both rules and algorithms should be assessed together. The methods and examples in this chapter can be used an initial framework to test the pricing outcomes.

6.2 FUTURE WORK ON RESERVE IMPROVEMENTS

As renewable penetration continues to grow, uncertainty and variability in the supply mix will increase. Although forecasting will improve over time, better reserve requirements will still be important for system reliability. The reserve study in Chapter 3 analyzed three improvements for reserve sizing, allocation, and activation. The improvements were chosen based on the perceived need in each area, but could be extended to include other concerns.

Instead of analyzing a simple reserve requirement, future work for the sizing phase could focus on some of the reserve proposals in the literature, such as probabilistic requirements. This could be especially useful since the comparison of a daily and seasonal requirement did not produce a significant difference in cost. Another possible improvement would be a separate coordination scenario where capacity on the lines was allocated for shared reserve. Both [92] and [114] found securing transmission line space to be an important aspect of trade, otherwise allowing increased trade between countries might not reduce costs. The simulations in Chapter 3 do not hold any capacity on the lines, possibly understating the impact of reserve coordination. Simulating the reservation of transmission capability could confirm the results from existing studies with additional quantitative analysis, and show greater impacts of coordinated reserve.

There are several areas where improved data or a different dataset can advance the analysis. For the contract-based analysis, gaining data on the quality, quantity, and price of existing contracts would enhance the comparison to the market-based case. The current comparison likely overestimates the costs, since the reserves were contracted based on a rule; however, there is no guarantee that the existing contracts would result in less costly solutions, since the details of those deals are private. Another area where additional data would enrich the simulations is the day-ahead wind forecast and actual wind generation data for multiple sites, along with contemporaneous load data. Newer data would reflect new wind sites and any changed inter-country correlations between wind farms. There is also more planned offshore wind than currently exists, and additional data might provide more insight into the value of offshore wind compared to onshore farms. Finally, simulations did not consider error in load forecasts, only wind forecasts. As discussed in Section 3.5.2, simulations with load forecast error would demonstrate the impacts of correlated wind and load on costs.

The comparison of alternative adjustments would also be improved with additional generation data, specifically, data on individual generators in countries surrounding the Netherlands. While generation data for all of Europe would be difficult to obtain, additional granularity on generators in Central Western Europe would provide more insight into the coordinated approaches. It might also change the reserve outcome; further correlation between countries might increase or decrease the requirement or increase congestion between countries. Finally, additional network data might provide further understanding to the amount of reserve that would be deliverable in real-time. Since the model only connects countries through a single line, actual delivery might increase or decrease depending on actual network congestion. A secondary real-time model with a granular network can answer questions on reserve deliverability and trade.

The three improvements studied in Chapter 3 were specific to the European market and relevant to the needs of the Netherlands. Future studies could use this same framework to analyze other regions or countries around the world. A similar study could be done in the Western U.S. to analyze the benefits of coordinating reserves in the day-ahead market between the California market and Western states. Such a study would complement the EIM in the real-time market and desire to form a west-wide ISO. Other regions can also use this framework to evaluate the benefits of making one or more improvements to existing reserve sizing, allocation, and activation.

6.3 FUTURE WORK ON EMISSIONS TRADING IN REGIONS WITH ASYMMETRICAL POLICIES

This chapter internalizes the cost of greenhouse gas emissions in markets with asymmetric policies and can benefit from further modeling options. Three approaches to accounting for and regulating carbon emissions associated with power imports were compared against two baseline (best and worst case) models, but other models are in development. Although many areas of the world have asymmetrical carbon or emissions policies, only recently has there been an increase in analytical and quantitative results. The modeling proposal in California has been modified several times since the inception of this dissertation chapter, and will likely continue to change in the future. As methods are proposed, the criteria developed in Section 4.5.2 can be used to evaluate and compare the new proposals against the existing second-best methods described in this chapter. If other regions in the U.S. decide to implement similar policies, these same criteria can be used to evaluate other systems. The results from Chapter 4 found tradeoffs between emissions and total cost for the second-best methods; however, the social cost of carbon method will always dominate. Although beyond the scope of optimization modeling, political analyses can evaluate what policies and political will would be necessary to implement a carbon tax through the U.S. (to produce the first-best, social cost of carbon result).

The project could also benefit from larger test cases. When initially investigating, it is important to begin with a small model so the results are easy to analyze and understand. However, as regions propose to implement these methods, it is important to recognize how the models scale-up. With more participants and a more complex network, behavior is likely to be qualitatively different; regions will sometimes be exporters and sometimes be importers. A West-wide model could provide much deeper insights into the actual emissions reductions and total system costs. A system with a more than one hour, so that a range of system load and renewable output conditions are represented, should also be studied, since it is average performance over a long period of time that matters the most, not performance in a single hour. A large system could also provide better understanding of the tax-at-the-border method, since the current simulation results jump between three options rather than a possibly smoother response when there are more plants and a diversity of load and renewable generation patterns over the year. A similar study can be performed in the mid-Atlantic region for RGGI and in Europe, possibly in collaboration with the existing COMPETES model used in Chapter 3.

6.4 FUTURE WORK ON COORDINATION AMONG BALANCING AREAS

Many countries and regions throughout the world have separately controlled but interconnected power systems, obliging system operators to coordinate the flow of electric power and trade. Because renewable energy is not always plentiful in areas with large populations, it can be important to improve trade between regions with abundant renewable generation and those with large load centers. Moving towards a majority renewable future, enabling trade can make all participants better off. Better off, of course, is relative to the aims of the organization or government in charge of operations.

The studies in Chapter 5 analyze trade across time and between regions with differing amounts of coordination. As renewable integration is increasing, there are many opportunities for model extensions. Foremost, would be extending the examples to more realistic test cases and testing against real system data. Many regions have existing hurdle rates (such as the WECC region), and use of actual hurdle rate data would produce results that mimic actual operations. Additionally, more than two regions can be simulated. Although the dynamics and constraints linking the systems would change, simulations using more than three balancing areas can demonstrate the challenges faced by highly interconnected grids. Larger systems might also allow simulation of game theoretic models to determine if a company that owns generators in multiple balancing areas can influence outcomes in either market.

The project can also extend modeling of trade to more complex mechanisms, some of which were discussed in Section 5.2. Both pseudo-tie lines and dynamic transfers can ease the integration of renewable energy from one balancing area to another. By explicitly modeling each mechanism, simulations can determine which would integrate a greater penetration of renewable energy or reduce costs. They can also be simulated in networks that use hurdle rates, to assess the relative benefits of using pseudo-tie lines and dynamic transfers.

Lastly, an empirical study of trade between regions using different mechanisms might identify causes and quantitative estimates of transaction costs that can be translated into hurdle rates. With the additional modeling, it can help address the question of the worth of hurdle rates. Is the mechanism too simple or does it adequately capture the complexity of inter-BA trade? When are hurdle rates useful as economic friction? Studies can assess current systems to determine if reductions to hurdle rates can increase gains from trade. As renewable integration grows, interconnection issues grow and become more important for ensuring the penetration is as high as possible.

In total, the four analyses provided in this thesis examine trade-offs between economic efficiency, simplicity of trade, integration of renewable energy, and emissions reductions. The dissertation contributes a framework for assessing market design improvements, and demonstrates that coordination between neighboring regions can increase economic efficiency. However, it is also important to remember the now infamous George Box quote, "All models are wrong but some are useful" [226]. Each chapter's models cannot exactly mimic real power system operation, but the insights provided can assist decision makers, researchers, system operators, and consumers in understanding electric market design options.

APPENDIX

Data used in Chapter 5 can be found in this Appendix. The tables show generator characteristics, line information, an hourly load profile, and bus/load levels.

Table 0-1 IEEE RTS Test Case Generator Data

Gen Number	Bus Location	Marginal Cost	Startup Cost	No Load Cost	Minimum Capacity
1	1	163.020	75.850	1138.680	15.8
2	1	163.020	75.850	1138.680	15.8
3	1	19.640	1060.880	130.630	15.2
4	1	19.640	1060.880	130.630	15.2
5	2	163.020	75.850	1138.680	15.8
6	2	163.020	75.850	1138.680	15.8
7	2	19.640	1060.880	130.630	15.2
8	2	19.640	1060.880	130.630	15.2
9	7	75.640	4754.400	839.450	25
10	7	75.640	4754.400	839.450	25
11	7	75.640	4754.400	839.450	25
12	13	74.750	6510.000	1159.930	68.95
13	13	74.750	6510.000	1159.930	68.95
14	13	74.750	6510.000	1159.930	68.95
15	14	0.000	0.000	0.000	0
16	15	94.740	571.200	72.680	2.4
17	15	94.740	571.200	72.680	2.4
18	15	94.740	571.200	72.680	2.4
19	15	94.740	571.200	72.680	2.4
20	15	94.740	571.200	72.680	2.4
21	15	15.460	1696.340	252.670	54.25
22	16	15.460	1696.340	252.670	54.25
23	18	5.460	2400.000	215.080	100
24	21	5.460	2400.000	215.080	100
25	22	0.000	0.000	0.000	0
26	22	0.000	0.000	0.000	0
27	22	0.000	0.000	0.000	0
28	22	0.000	0.000	0.000	0

Gen Number	Bus Location	Marginal Cost	Startup Cost	No Load Cost	Minimum Capacity
29	22	0.000	0.000	0.000	0
30	22	0.000	0.000	0.000	0
31	23	15.460	1696.340	252.670	54.25
32	23	15.460	1696.340	252.670	54.25
33	23	15.890	7953.040	358.230	140
34	25	163.020	75.850	1138.680	15.8
35	25	163.020	75.850	1138.680	15.8
36	25	19.640	1060.880	130.630	15.2
37	25	19.640	1060.880	130.630	15.2
38	26	163.020	75.850	1138.680	15.8
39	26	163.020	75.850	1138.680	15.8
40	26	19.640	1060.880	130.630	15.2
41	26	19.640	1060.880	130.630	15.2
42	31	75.640	4754.400	839.450	25
43	31	75.640	4754.400	839.450	25
44	31	75.640	4754.400	839.450	25
45	37	74.750	6510.000	1159.930	68.95
46	37	74.750	6510.000	1159.930	68.95
47	37	74.750	6510.000	1159.930	68.95
48	38	0.000	0.000	0.000	0
49	39	94.740	571.200	72.680	2.4
50	39	94.740	571.200	72.680	2.4
51	39	94.740	571.200	72.680	2.4
52	39	94.740	571.200	72.680	2.4
53	39	94.740	571.200	72.680	2.4
54	39	15.460	1696.340	252.670	54.25
55	40	15.460	1696.340	252.670	54.25
56	42	5.460	2400.000	215.080	100
57	45	5.460	2400.000	215.080	100
58	46	0.000	0.000	0.000	0
59	46	0.000	0.000	0.000	0
60	46	0.000	0.000	0.000	0
61	46	0.000	0.000	0.000	0
62	46	0.000	0.000	0.000	0
63	46	0.000	0.000	0.000	0
64	47	15.460	1696.340	252.670	54.25
65	47	15.460	1696.340	252.670	54.25
66	47	15.890	7953.040	358.230	140
67	49	163.020	75.850	1138.680	15.8
68	49	163.020	75.850	1138.680	15.8
69	49	19.640	1060.880	130.630	15.2
70	49	19.640	1060.880	130.630	15.2
71	50	163.020	75.850	1138.680	15.8
72	50	163.020	75.850	1138.680	15.8
73	50	19.640	1060.880	130.630	15.2
74	50	19.640	1060.880	130.630	15.2
75	55	75.640	4754.400	839.450	25
76	55	75.640	4754.400	839.450	25
77	55	75.640	4754.400	839.450	25

Gen Number	Bus Location	Marginal Cost	Startup Cost	No Load Cost	Minimum Capacity
78	61	74.750	6510.000	1159.930	68.95
79	61	74.750	6510.000	1159.930	68.95
80	61	74.750	6510.000	1159.930	68.95
81	62	0.000	0.000	0.000	0
82	63	94.740	571.200	72.680	2.4
83	63	94.740	571.200	72.680	2.4
84	63	94.740	571.200	72.680	2.4
85	63	94.740	571.200	72.680	2.4
86	63	94.740	571.200	72.680	2.4
87	63	15.460	1696.340	252.670	54.25
88	64	15.460	1696.340	252.670	54.25
89	66	5.460	2400.000	215.080	100
90	69	5.460	2400.000	215.080	100
91	70	0.000	0.000	0.000	0
92	70	0.000	0.000	0.000	0
93	70	0.000	0.000	0.000	0
94	70	0.000	0.000	0.000	0
95	70	0.000	0.000	0.000	0
96	70	0.000	0.000	0.000	0
97	71	15.460	1696.340	252.670	54.25
98	71	15.460	1696.340	252.670	54.25
99	71	15.890	7953.040	358.230	140

Table 0-2 IEEE RTS Test Case Generator Data

Generator Number	Ramp Rate	10 Min. Ramp Rate	Down Time	Up Time	Fuel Type
1	20	20	1	1	Oil/CT
2	20	20	1	1	Oil/CT
3	76	20	4	8	Coal/Steam
4	76	20	4	8	Coal/Steam
5	20	20	1	1	Oil/CT
6	20	20	1	1	Oil/CT
7	76	20	4	8	Coal/Steam
8	76	20	4	8	Coal/Steam
9	100	70	8	8	Oil/Steam
10	100	70	8	8	Oil/Steam
11	100	70	8	8	Oil/Steam
12	180	30	10	12	Oil/Steam
13	180	30	10	12	Oil/Steam
14	180	30	10	12	Oil/Steam
15	0	0	1	1	N/A
16	12	10	2	4	Oil/Steam
17	12	10	2	4	Oil/Steam
18	12	10	2	4	Oil/Steam
19	12	10	2	4	Oil/Steam
20	12	10	2	4	Oil/Steam
21	155	30	8	8	Coal/Steam

Generator Number	Ramp Rate	10 Min. Ramp Rate	Down Time	Up Time	Fuel Type
22	155	30	8	8	Coal/Steam
23	400	200	24	24	Nuclear
24	400	200	24	24	Nuclear
25	50	50	1	1	Hydro
26	50	50	1	1	Hydro
27	50	50	1	1	Hydro
28	50	50	1	1	Hydro
29	50	50	1	1	Hydro
30	50	50	1	1	Hydro
31	155	30	8	8	Coal/Steam
32	155	30	8	8	Coal/Steam
33	240	40	24	24	Coal/Steam
34	20	20	1	1	Oil/CT
35	20	20	1	1	Oil/CT
36	76	20	4	8	Coal/Steam
37	76	20	4	8	Coal/Steam
38	20	20	1	1	Oil/CT
39	20	20	1	1	Oil/CT
40	76	20	4	8	Coal/Steam
41	76	20	4	8	Coal/Steam
42	100	70	8	8	Oil/Steam
43	100	70	8	8	Oil/Steam
44	100	70	8	8	Oil/Steam
45	180	30	10	12	Oil/Steam
46	180	30	10	12	Oil/Steam
47	180	30	10	12	Oil/Steam
48	0.000	0	1	1	N/A
49	12	10	2	4	Oil/Steam
50	12	10	2	4	Oil/Steam
51	12	10	2	4	Oil/Steam
52	12	10	2	4	Oil/Steam
53	12	10	2	4	Oil/Steam
54	155	30	8	8	Coal/Steam
55	155	30	8	8	Coal/Steam
56	400	200	24	24	Nuclear
57	400	200	24	24	Nuclear
58	50	50	1	1	Hydro
59	50	50	1	1	Hydro
60	50	50	1	1	Hydro
61	50	50	1	1	Hydro
62	50	50	1	1	Hydro
63	50	50	1	1	Hydro
64	155	30	8	8	Coal/Steam
65	155	30	8	8	Coal/Steam
66	240	40	24	24	Coal/Steam
67	20	20	1	1	Oil/CT
68	20	20	1	1	Oil/CT
69	76	20	4	8	Coal/Steam
70	76	20	4	8	Coal/Steam

Generator Number	Ramp Rate	10 Min. Ramp Rate	Down Time	Up Time	Fuel Type
71	20	20	1	1	Oil/CT
72	20	20	1	1	Oil/CT
73	76	20	4	8	Coal/Steam
74	76	20	4	8	Coal/Steam
75	100	70	8	8	Oil/Steam
76	100	70	8	8	Oil/Steam
77	100	70	8	8	Oil/Steam
78	180	30	10	12	Oil/Steam
79	180	30	10	12	Oil/Steam
80	180	30	10	12	Oil/Steam
81	0	0	1	1	N/A
82	12	10	2	4	Oil/Steam
83	12	10	2	4	Oil/Steam
84	12	10	2	4	Oil/Steam
85	12	10	2	4	Oil/Steam
86	12	10	2	4	Oil/Steam
87	155	30	8	8	Coal/Steam
88	155	30	8	8	Coal/Steam
89	400	200	24	24	Nuclear
90	400	200	24	24	Nuclear
91	50	50	1	1	Hydro
92	50	50	1	1	Hydro
93	50	50	1	1	Hydro
94	50	50	1	1	Hydro
95	50	50	1	1	Hydro
96	50	50	1	1	Hydro
97	155	30	8	8	Coal/Steam
98	155	30	8	8	Coal/Steam
99	240	40	24	24	Coal/Steam

Table 0-3 IEEE RTS Test Case Line Data

Bus Number	From	To	Susceptance	Line Capacity
1	1	2	-7142.8571	175
2	1	3	-473.9336	175
3	1	5	-1176.4706	175
4	2	4	-787.4016	175
5	2	6	-520.8333	175
6	3	9	-840.3361	175
7	3	24	-1190.4762	400
8	4	9	-961.5385	175
9	5	10	-1136.3636	175
10	6	10	-1639.3443	175
11	7	8	-1639.3443	175
12	7	27	-621.1180	175
13	8	9	-606.0606	175
14	8	10	-606.0606	175

Bus Number	From	To	Susceptance	Line Capacity
15	9	11	-1190.4762	400
16	9	12	-1190.4762	400
17	10	11	-1190.4762	400
18	10	12	-1190.4762	400
19	11	14	-2380.9524	500
20	12	13	-2083.3333	500
21	12	23	-1030.9278	500
22	13	23	-1149.4253	500
23	13	39	-1333.3333	500
24	14	16	-1694.9153	350
25	15	16	-5882.3529	500
26	15	21	-2040.8163	500
27	15	21	-2040.8163	500
28	15	24	-1923.0769	500
29	16	17	-3846.1538	500
30	16	19	-4347.8261	500
31	17	18	-7142.8571	500
32	17	22	-952.3810	500
33	18	21	-3846.1538	500
34	18	21	-3846.1538	500
35	19	20	-2500.0000	500
36	19	20	-2500.0000	500
37	20	23	-4545.4545	500
38	20	23	-4545.4545	500
39	21	22	-1470.5882	500
40	23	41	-1351.3514	500
41	25	26	-7142.8571	175
42	25	27	-473.9336	175
43	25	29	-1176.4706	175
44	26	28	-787.4016	175
45	26	30	-520.8333	175
46	27	33	-840.3361	175
47	27	48	-1190.4762	400
48	28	33	-961.5385	175
49	29	34	-1136.3636	175
50	30	34	-1639.3443	175
51	31	32	-1639.3443	175
52	32	33	-606.0606	175
53	32	34	-606.0606	175
54	33	35	-1190.4762	400
55	33	36	-1190.4762	400
56	34	35	-1190.4762	400
57	34	36	-1190.4762	400
58	35	38	-2380.9524	500
59	36	37	-2083.3333	500
60	36	47	-1030.9278	500
61	37	47	-1149.4253	500
62	38	40	-1694.9153	350
63	39	40	-5882.3529	500

Bus Number	From	To	Susceptance	Line Capacity
64	39	45	-2040.8163	500
65	39	45	-2040.8163	500
66	39	48	-1923.0769	500
67	40	41	-3846.1538	500
68	40	43	-4347.8261	500
69	41	42	-7142.8571	500
70	41	46	-952.3810	500
71	42	45	-3846.1538	500
72	42	45	-3846.1538	500
73	43	44	-2500.0000	500
74	43	44	-2500.0000	500
75	44	47	-4545.4545	500
76	44	47	-4545.4545	500
77	45	46	-1470.5882	500
78	49	50	-7142.8571	175
79	49	51	-473.9336	175
80	49	53	-1176.4706	175
81	50	52	-787.4016	175
82	50	54	-520.8333	175
83	51	57	-840.3361	175
84	51	72	-1190.4762	400
85	52	57	-961.5385	175
86	53	58	-1136.3636	175
87	54	58	-1639.3443	175
88	55	56	-1639.3443	175
89	56	57	-606.0606	175
90	56	58	-606.0606	175
91	57	59	-1190.4762	400
92	57	60	-1190.4762	400
93	58	59	-1190.4762	400
94	58	60	-1190.4762	400
95	59	62	-2380.9524	500
96	60	61	-2083.3333	500
97	60	71	-1030.9278	500
98	61	71	-1149.4253	500
99	62	64	-1694.9153	350
100	63	64	-5882.3529	500
101	63	69	-2040.8163	500
102	63	69	-2040.8163	500
103	63	72	-1923.0769	500
104	64	65	-3846.1538	500
105	64	67	-4347.8261	500
106	65	66	-7142.8571	500
107	65	70	-952.3810	500
108	66	69	-3846.1538	500
109	66	69	-3846.1538	500
110	67	68	-2500.0000	500
111	67	68	-2500.0000	500
112	68	71	-4545.4545	500

Bus Number	From	To	Susceptance	Line Capacity
113	68	71	-4545.4545	500
114	69	70	-1470.5882	500
115	73	21	-1030.9278	500
116	66	47	-961.5385	500
117	71	73	-11111.1111	722

Table 0-4 IEEE RTS Test Case Bus Data

Bus Number	Peak Bus Demand
1	108
2	97
3	180
4	74
5	71
6	136
7	125
8	171
9	175
10	195
11	0
12	0
13	745
14	80.4
15	131.4
16	100
17	0
18	333
19	75.1
20	53.1
21	0
22	0
23	0
24	0
25	108
26	97
27	180
28	74
29	71
30	136
31	125
32	171
33	175
34	195
35	0
36	0
37	745
38	80.4

Bus Number	Peak Bus Demand
39	131.4
40	100
41	0
42	333
43	75.1
44	53.1
45	0
46	0
47	0
48	0
49	108
50	97
51	180
52	74
53	71
54	136
55	125
56	171
57	175
58	195
59	0
60	0
61	745
62	80.4
63	131.4
64	100
65	0
66	333
67	75.1
68	53.1
69	0
70	0
71	0
72	0
73	0

Table 0-5 IEEE RTS Test Case Hourly Data

Hour	Percent	Peak Hourly Demand
1	67	1909.5
2	63	1795.5
3	60	1710
4	59	1681.5
5	59	1681.5
6	60	1710
7	74	2109
8	86	2451
9	95	2707.5
10	96	2736
11	96	2736
12	95	2707.5
13	95	2707.5
14	95	2707.5
15	93	2650.5
16	94	2679
17	99	2821.5
18	100	2850
19	100	2850
20	96	2736
21	91	2593.5
22	83	2365.5
23	73	2080.5
24	63	1795.5

REFERENCES

[1] U.S. Department of Energy, “The smart grid: an introduction,” U.S. Department of Energy, Washington D.C., 2010.

[2] S. Borenstein and J. Bushnell, “The U.S. electricity industry after 20 years of restructuring,” *Annu. Rev. Econom.*, vol. 7, pp. 437–463, 2015.

[3] W. M. Warwick, “A primer on electric utilities, deregulation, and restructuring of us electricity markets,” PNNL for the US. DOE, Office of Energy Efficiency and Renewable Energy, Washington D.C., 2002.

[4] Divison of Energy Market Oversight, “Energy primer: a handbook of energy market basics,” FERC, Washington D.C., 2015.

[5] FERC Docket No. AD14-14-000, “Price formation in energy and ancillary services markets operated by regional transmission organizations and independent system operators: order directing reports,” 2015. [Online]. Available: elibrary.ferc.gov/idmws/file_list.asp?document_id=14400898.

[6] FERC, “Ferc to investigate pricing of fast-start resources by three grid operators,” *News Release*. [Online]. Available: www.ferc.gov/media/news-releases/2017/2017-4/12-21-17-E-2.asp#.WpwlLpPwbMV.

[7] “Seams issues high priority items,” NYISO Committee Proposal, Rensselaer, NY, 2001.

[8] CAISO, “Fifth replacement tariff, section 29 energy imbalance market,” 2018. [Online].

Available:

www.caiso.com/Documents/Section29_EnergyImbalanceMarket_asof_Feb15_2018.pdf.

[9] H. Holttinen, P. Meibom, A. Orths, F. van Hulle, B. Lange, M. O'Malley, J. Pierik, B. Ummels, J. O. Tande, A. Estanqueiro, M. Matos, J. Ricardo, E. Gomez, L. Soder, G. Strbac, A. Shakoor, J. C. Smith, M. Milligan, and E. Ela, "IEA wind task 25 design and operation of power systems with large amounts of wind power," 2009.

[10] "Issues affecting renewable energy integration," *NREL Transmission Grid Integration*, 2014. [Online]. Available: www.nrel.gov/electricity/transmission/issues.html.

[11] Y. Chen, A. L. Liu, and B. F. Hobbs, "Economic and emissions implications of load-based, source-based, and first-seller emissions trading programs under California AB32," *Oper. Res.*, vol. 59, no. 3, pp. 696–712, 2011.

[12] W. Dean, "Interactions among market mechanisms for reducing greenhouse gas emissions in California," *Electr. J.*, vol. 29, no. 8, pp. 17–22, 2016.

[13] F. C. Schweppe, M. C. Caramanis, R. D. Tabors, and R. E. Bohn, *Spot Pricing of Electricity*. Boston, MA: Kluwer Academic Publisher, 1988.

[14] U. C. § 824d, *Rates and charges; schedules; suspension of new rates; automatic adjustment clauses*. Washington D.C.

[15] H. R. Varian, *Intermediate Microeconomics: A Modern Approach*, 8th ed. New York: W. W. Norton & Company, 2014.

[16] H. P. Williams, "Duality in mathematics and linear and integer programming," *J. Optim. Theory Appl.*, vol. 90, no. 2, pp. 257–278, 1996.

[17] D. Bertsimas and J. N. Tsitsiklis, *Introduction to Linear Optimization*. Belmont, MA: Athena Scientific and Dynamic Ideas, 1997.

[18] L. A. Wolsey, "Integer programming duality: price functions and sensitivity analysis," *Math. Program.*, vol. 20, no. 1, pp. 173–195, 1981.

[19] "Directive of the european parliament and of the council on the promotion of the use of

energy from renewable sources (recast)," European Commission, Brussels, Belgium, 2016.

[20] S. C. Pryor, R. J. Barthelmie, and J. T. Schoof, "Inter-annual variability of wind indices across Europe," *Wind Energy*, vol. 9, no. 1–2, pp. 27–38, 2006.

[21] R. K. Pachauri and L. A. Meyer, "Climate change 2014: synthesis report, contribution of working groups I, II and III to the fifth assessment report of the Intergovernmental Panel on Climate Change," IPCC, 2014.

[22] "Chapter 12. environment," *Monthly Energy Review*, EIA, 2018, [Online]. Available: www.eia.gov/totalenergy/data/monthly/index.php#environment.

[23] S. A. Newell, R. Lueken, J. Weiss, K. Spees, P. Donohoo-Vallett, and T. Lee, "Pricing carbon into NYISO's wholesale energy market to support New York's decarbonization goals," The Brattle Group, Prepared for NYISO, 2017.

[24] H. Bin Zhang, H. C. Dai, H. X. Lai, and W. T. Wang, "U.S. withdrawal from the Paris agreement: reasons, impacts, and China's response," *Adv. Clim. Chang. Res.*, vol. 8, no. 4, pp. 220–225, 2017.

[25] PJM, "Advancing zero emissions objectives through PJM's energy markets: a review of carbon-pricing frameworks," 2017. [Online]. Available: www.pjm.com/~/media/library/reports-notices/special-reports/20170502-advancing-zero-emission-objectives-through-pjms-energy-markets.ashx.

[26] California Assembly, "California global warming solutions act," *California State Law*, 2006. [Online]. Available: www.arb.ca.gov/cc/ab32/ab32.htm.

[27] W. W. Hogan, "An efficient Western Energy Imbalance market with conflicting carbon policies," *Electr. J.*, vol. 30, no. 10, pp. 8–15, 2017.

[28] M. H. Babiker, "Climate change policy , market structure , and carbon leakage," vol. 65, pp. 421–445, 2005.

[29] A. Antimiani, V. Costantini, C. Martini, L. Salvatici, and M. Cristina, "Assessing

alternative solutions to carbon leakage," *Energy Econ.*, vol. 36, pp. 299–311, 2013.

[30] E. Sauma, "The impact of transmission constraints on the emissions leakage," *Energy Policy*, vol. 51, pp. 164–171, 2012.

[31] Y. Chen, "Does a regional greenhouse gas policy make sense? A case study of carbon leakage and emissions spillover," *Energy Econ.*, vol. 31, no. 5, pp. 667–675, 2009.

[32] CAISO, "Fifth replacement tariff, section 33 hour-ahead scheduling process," 2013. [Online]. Available: www.caiso.com/Documents/Section33_HourAheadSchedulingProcess_HASP_Jul1_2013.pdf.

[33] C. Özden-Schilling, "Economy electric," *Cult. Anthropol.*, vol. 30, no. 4, pp. 578–588, 2015.

[34] M. Milligan, B. Frew, K. Clark, and A. Bloom, "Marginal cost pricing in a world without perfect competition: implications for electricity markets with high shares of low marginal cost resources," NREL/TP-6A20-69076, Golden, CO, 2017.

[35] FERC, "Order directing report, Docket No. AD 14-14-000," 2015. [Online]. Available: elibrary.ferc.gov/idmws/file_list.asp?document_id=14400898.

[36] PJM Interconnection, "PJM cold snap performance Dec. 28, 2017 to Jan. 7, 2018," 2018. [Online]. Available: www.pjm.com/-/media/library/reports-notices/weather-related/20180226-january-2018-cold-weather-event-report.ashx.

[37] FERC Docket No. EL08-48-000, "Braintree Electric Light Department v. ISO New England Inc.," 2008. [Online]. Available: www.ferc.gov/whats-new/comm-meet/2008/071708/E-30.pdf.

[38] FERC Docket No. EL14-34-000, "Public service commission of Wisconsin v. Midcontinent Independent System Operator, Inc," 2014. [Online]. Available: www.ferc.gov/CalendarFiles/20140729142046-ER14-1242-000.pdf.

[39] R. O'Neill, A. Castillo, B. Eldridge, and R. B. Hytowitz, "Dual Pricing Algorithm in ISO

markets," *IEEE Trans. Power Syst.*, vol. 32, no. 4. pp. 3308–3310, 2017.

[40] H. Chao and H. G. Huntington, *Designing Competitive Electricity Markets*, 1st ed. New York: Springer Science+Business Media, 1998.

[41] W. Hogan, "Competitive electricity market design: a wholesale primer," John F. Kennedy School of Government, Harvard University, Cambridge, MA, 1998.

[42] P. Cramton, "Electricity market design: the good, the bad, and the ugly," *Syst. Sci. 2003. Proc. 36th Annu. Hawaii Int. Conf.*, p. 8–pp, 2003.

[43] R. P. O'Neill, P. M. Sotkiewicz, B. F. Hobbs, M. H. Rothkopf, and W. R. Stewart Jr., "Efficient market-clearing prices in markets with nonconvexities," *Eur. J. Oper. Res.*, vol. 164, no. 1, pp. 269–285, Jul. 2005.

[44] M. Van Vyve, "Linear prices for non-convex electricity markets: models and algorithms," 2011.

[45] PJM, "Problem statement day ahead surplus congestion and FTR auction revenue surplus funds," 2017. [Online]. Available: www.pjm.com/~/media/committees-groups/committees/mic/20170412/20170412-item-09a-ftr-surplus-funds-problem-statement.ashx.

[46] T. Kristiansen, "Markets for financial information," *Energy Stud. Rev.*, vol. 13, no. 1, pp. 25–74, 2004.

[47] I. Herrero, P. Rodilla, and C. Batlle, "Electricity market-clearing prices and investment incentives: the role of pricing rules," *Energy Econ.*, vol. 47, pp. 42–51, 2015.

[48] M. Caramanis, R. Bohn, and F. Schweppe, "Optimal spot pricing: practice and theory," *IEEE Trans. Power Appar. Syst.*, vol. PAS-101, no. 9, pp. 3234–3245, 1982.

[49] G. Liberopoulos and P. Andrianesis, "Critical review of pricing schemes in markets with non-convex costs," *Oper. Res.*, vol. 64, no. 1, pp. 17–31, 2016.

[50] M. Bjørndal and K. Jörnsten, "Equilibrium prices supported by dual price functions in markets with non-convexities," *Eur. J. Oper. Res.*, vol. 190, no. 3, pp. 768–789, Nov. 2008.

[51] Y. M. Al-Abdullah, M. Abdi-Khorsand, and K. W. Hedman, "The role of out-of-market corrections in day-ahead scheduling," *IEEE Trans. Power Syst.*, vol. 30, no. 4, pp. 1937–1946, 2015.

[52] F. Ramsey, "A contribution to the theory of taxation," *Econ. J.*, vol. 37, pp. 47–61, 1927.

[53] M. Boiteux, "Sur la gestion des monopoles publics astreints a l'équilibre budgétaire," *Econometrica*, vol. 24, no. 1, pp. 22–40, 1956.

[54] W. W. Hogan and B. J. Ring, "On minimum-uplift pricing for electricity markets," Cambridge, MA, 2003.

[55] P. R. Gribik, W. W. Hogan, and S. L. Pope, "Market-clearing electricity prices and energy uplift," Cambridge, MA, 2007.

[56] C. Wang, P. B. Luh, P. Gribik, L. Zhang, and T. Peng, "The subgradient-simplex based cutting plane method for convex hull pricing," *IEEE PES Gen. Meet. PES 2010*, pp. 1–8, 2010.

[57] C. Wang, T. Peng, P. B. Luh, P. Gribik, and L. Zhang, "The subgradient simplex cutting plane method for extended locational marginal prices," *IEEE Trans. Power Syst.*, vol. 28, no. 3, pp. 2758–2767, 2013.

[58] "Extended LMP," *Midcontinent Independent System Operator, Inc.*, 2016. [Online]. Available: www.misoenergy.org/WhatWeDo/MarketEnhancements/Pages/ELMP.aspx.

[59] A. L. Motto and F. D. Galiana, "Equilibrium of auction markets with unit commitment: the need for augmented pricing," *IEEE Trans. Power Syst.*, vol. 17, no. 3. pp. 798–805, 2002.

[60] F. D. Galiana, A. L. Motto, and F. Bouffard, "Reconciling social welfare, agent profits, and consumer payments in electricity pools," *IEEE Trans. Power Syst.*, vol. 18, no. 2. pp. 452–459, 2003.

[61] V. Araoz and K. Jörnsten, "Semi-Lagrangean approach for price discovery in markets with non-convexities," *Eur. J. Oper. Res.*, vol. 214, no. 2, pp. 411–417, Oct. 2011.

[62] C. Ruiz, A. J. Conejo, and S. A. Gabriel, "Pricing non-convexities in an electricity pool," *IEEE Trans. Power Syst.*, vol. 27, no. 3. pp. 1334–1342, 2012.

[63] P. Andrianesis, G. Liberopoulos, G. Kozanidis, and A. D. Papalexopoulos, "Recovery mechanisms in day-ahead electricity markets with non-convexities - part I: design and evaluation methodology," *IEEE Trans. Power Syst.*, vol. 28, no. 2. pp. 960–968, 2013.

[64] R. Sioshansi, R. O'Neill, and S. S. Oren, "Economic consequences of alternative solution methods for centralized unit commitment in day-ahead electricity markets," *IEEE Trans. Power Syst.*, vol. 23, no. 2. pp. 344–352, 2008.

[65] D. Huppmann and S. Siddiqui, "An exact solution method for binary equilibrium problems with compensation and the power market uplift problem," *Eur. J. Oper. Res.*, vol. 266, no. 2, pp. 622–638, 2018.

[66] H. P. Young, *Cost Allocation: Methods, Principles, Applications*, 1st ed. Amsterdam: Elseviers Science Publishers B.V., 1985.

[67] H. Rudnick, R. Palma, and J. E. Fernandez, "Marginal pricing and supplement cost allocation in transmission open access," *IEEE Trans. Power Syst.*, vol. 10, no. 2. pp. 1125–1132, 1995.

[68] F. J. Rubio-Oderiz and I. J. Perez-Arriaga, "Marginal pricing of transmission services: a comparative analysis of network cost allocation methods," *IEEE Trans. Power Syst.*, vol. 15, no. 1. pp. 448–454, 2000.

[69] S. J. Rassenti, V. L. Smith, and B. J. Wilson, "Controlling market power and price spikes in electricity networks: demand-side bidding.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 100, no. 5, pp. 2998–3003, 2003.

[70] C. L. Su, "Optimal demand-side participation in day-ahead electricity markets," The University of Manchester, 2007.

[71] C. Su and D. Kirschen, "Quantifying the effect of demand response on electricity markets," *Power Syst. IEEE Trans.*, vol. 24, no. 3, pp. 1199–1207, 2009.

[72] J. Torriti, M. G. Hassan, and M. Leach, "Demand response experience in Europe: policies, programmes and implementation," *Energy*, vol. 35, no. 4, pp. 1575–1583, 2010.

[73] H. Chao, "Demand response in wholesale electricity markets: the choice of customer baseline," *J. Regul. Econ.*, vol. 39, no. 1, pp. 68–88, 2011.

[74] B. Foster, D. Burns, J. Grove, D. Kathan, M. P. Lee, S. Peirovi, and C. Schilling, "2017 assessment of demand response and advanced metering," FERC, Washington D.C., 2017.

[75] G. Murtaugh and Department of Market Monitoring, "2016 annual report on market issues and performance," CAISO, Folsom, CA, 2017.

[76] NYISO, "Demand response programs," 2016. [Online]. Available: www.nyiso.com/public/markets_operations/market_data/demand_response/index.jsp.

[77] R. O'Neill, A. Castillo, B. Eldridge, and R. B. Hytowitz, "Dual pricing algorithm in ISO markets," *IEEE Trans. Power Syst.*, vol. PP, no. 99, p. 1, 2016.

[78] NYISO, "Accounting and billing manual," *Manual 14*, 2016. [Online]. Available: www.nyiso.com/public/webdocs/markets_operations/documents/Manuals_and_Guides/Manuals/Administrative/acctbillmnl.pdf.

[79] CAISO, "California Independent System Operator Corporation fifth replacement tariff, section 11.23." [Online]. Available: www.caiso.com/Documents/ConformedTariff_asof_Jul10_2017.pdf.

[80] CAISO, "Attachment D – benchmarking against NYISO, PJM, and ISO-NE." [Online]. Available: www.caiso.com/Documents/AttachmentD-BenchmarkingAgainstNYISO_PJMandISO-NE.pdf.

[81] ISO-NE, "FAQs: net commitment-period compensation." [Online]. Available: www.iso-ne.com/participate/support/faq/ncpc-rmr.

[82] Potomac Economics, "2016 state of the market report for the MISO electricity markets." [Online]. Available: www.potomaceconomics.com/wp-content/uploads/2017/06/2016-

SOM_Report_Final_6-30-17.pdf.

[83] P. Gribik and L. Zhang, "Extended Locational Marginal Pricing (Convex Hull Pricing)," in *FERC Technical Conference on Unit Commitment Software*, 2010, pp. 1–16.

[84] D. A. Schiro, T. Zheng, F. Zhao, and E. Litvinov, "Convex hull pricing in electricity markets: formulation, analysis, and implementation challenges," *IEEE Trans. Power Syst.*, vol. 31, no. 5, pp. 4068–4075, 2016.

[85] H. Scarf, "The allocation of resources in the presence of indivisibilities," *J. Econ. Perspect.*, vol. 8, no. 4, pp. 111–128, 1994.

[86] C. Grigg, "Power systems test case archive: reliability test system (rts) - 1996," 1996. .

[87] S. F. Tierney, T. Schatzki, and R. Mukerji, "Uniform-pricing versus pay-as-bid in wholesale electricity markets: does it make a difference?," Analyisis Group and NYISO, 2008.

[88] E. Ela, M. Milligan, and B. Kirby, "Operating reserves and variable generation," National Renewable Energy Laboratory TP-5500-51978, Golden, CO, 2011.

[89] T. Brijs, C. De Jonghe, B. F. Hobbs, and R. Belmans, "Interactions between the design of short-term electricity markets in the CWE region and power system flexibility," *Appl. Energy*, vol. 195, pp. 36–51, 2017.

[90] NERC, "Long-term reliability assessments," *Reliability Assessment and Performance Analysis*, 2018. [Online]. Available: www.nerc.com/pa/RAPA/ra/Pages/default.aspx.

[91] M. D. Bartos and M. V. Chester, "Impacts of climate change on electric power supply in the Western United States," *Nat. Clim. Chang.*, vol. 5, no. 8, pp. 748–752, 2015.

[92] K. Van den Bergh, R. B. Hytowitz, K. Bruninx, E. Delarue, W. Dhaeseleer, and B. F. Hobbs, "Benefits of coordinating sizing, allocation and activation of reserves among market zones," *Electr. Power Syst. Res.*, vol. 143, pp. 140–148, 2017.

[93] M. Milligan, P. Donohoo, D. Lew, E. Ela, B. Kirby, H. Holttinen, E. Lannoye, D. Flynn, M. O'Malley, N. Miller, P. B. Eriksen, A. Gøttig, B. Rawn, M. Gibescu, E. G. Lázaro, A.

Robitaille, and I. Kamwa, “Operating reserves and wind power integration: an international comparison,” in *The 9th Annual International Workshop on Large-Scale Integration of Wind Power into Power Systems as well as on Transmission Networks for Offshore Wind Power Plants Conference*, 2010, no. October 2010.

[94] P. González, J. Villar, C. A. Díaz, and F. A. Campos, “Joint energy and reserve markets: current implementations and modeling trends,” *Electr. Power Syst. Res.*, vol. 109, pp. 101–111, Apr. 2014.

[95] E. Ela, B. Kirby, E. Lannoye, M. Milligan, D. Flynn, B. Zavadil, and M. O’Malley, “Operating reserve determination evolution in wind power integration studies,” *Power Energy Soc. Gen. Meet. 2010*, no. March 2011, pp. 1–8, 2010.

[96] TenneT, “Primary reserve,” *System Data Preparation*, 2017. [Online]. Available: www.tennet.org/english/operational_management/system_data_preparation/primary_reserve.aspx.

[97] TenneT, “Product information aFRR (regulating power),” 2016. [Online]. Available: www.tennet.eu/fileadmin/user_upload/Company/Publications/Technical_Publications/Dutch/Product_information_aFRR_-regulating_power-_25-08-2016.pdf.

[98] J. Villar, R. Bessa, and M. Matos, “Flexibility products and markets: literature review,” *Electr. Power Syst. Res.*, vol. 154, pp. 329–340, 2018.

[99] M. A. Ortega-Vazquez and D. S. Kirschen, “Optimizing the spinning reserve requirements using a cost/benefit analysis,” *IEEE Trans. Power Syst.*, vol. 22, no. 1, pp. 24–33, 2007.

[100] M. A. Ortega-Vazquez and D. S. Kirschen, “Estimating the spinning reserve requirements in systems with significant wind power generation penetration,” *IEEE Trans. Power Syst.*, vol. 24, no. 1. pp. 114–124, 2009.

[101] R. Doherty and M. O’Malley, “A new approach to quantify reserve demand in systems with significant installed wind capacity,” *IEEE Trans. Power Syst.*, vol. 20, no. 2, pp. 587–

595, 2005.

[102] J. M. Morales, A. J. Conejo, and J. Pérez-Ruiz, "Economic valuation of reserves in power systems with high penetration of wind power," *IEEE Trans. Power Syst.*, vol. 24, no. 2, pp. 900–910, 2009.

[103] K. Bruninx and E. Delarue, "Endogenous probabilistic reserve sizing and allocation in unit commitment models: cost effective, reliable, and fast," *IEEE Trans. Power Syst.*, vol. 32, no. 4, pp. 2593–2603, 2017.

[104] J. F. Restrepo and F. D. Galiana, "Effects of wind power on day-ahead reserve schedule," in *1st IEEE-PES/IAS Conference on Sustainable Alternative Energy*, 2009, no. 2, pp. 1–4.

[105] I. Herrero, P. Rodilla, and C. Batlle, "Enhancing intraday price signals in U.S. ISO markets for a better integration of variable energy resources," Cambridge, MA, 2016.

[106] ENTSO-E WGAS, "Survey on ancillary services procurement, balancing market design 2014," Brussels, 2015.

[107] R. Scharff, J. Egerer, and L. Söder, "A description of the operative decision-making process of a power generating company on the Nordic electricity market," *Energy Syst.*, vol. 5, no. 2, pp. 349–369, 2014.

[108] F. Tanrisever, K. Derinkuyu, and G. Jongen, "Organization and functioning of liberalized electricity markets : an overview of the Dutch market," *Renew. Sustain. Energy Rev.*, vol. 51, no. 11, pp. 1363–1374, 2015.

[109] G. Brunekreeft, "Empirics of intraday and real-time markets in Europe: the Netherlands," 2015.

[110] B. C. Ummels, M. Gibescu, E. Pelgrum, and W. L. Kling, "System integration of large-scale wind power in the Netherlands," in *IEEE Power Engineering Society General Meeting*, 2006, pp. 1–8.

[111] D. Newbery, M. Pollitt, R. Ritz, and W. Strielkowski, "Market design for a high-

renewables european electricity system," *Cambridge Work. Pap. Econ.*, vol. 1711, 2017.

[112] G. L. Doorman and R. Van Der Veen, "An analysis of design options for markets for cross-border balancing of electricity," *Util. Policy*, vol. 27, pp. 39–48, 2013.

[113] H. Farahmand and G. L. Doorman, "Balancing market integration in the Northern European continent," *Appl. Energy*, vol. 96, pp. 316–326, Aug. 2012.

[114] Y. Gebrekiros and G. Doorman, "Optimal transmission capacity allocation for cross-border exchange of frequency restoration reserves (FRR)," *Proc. - 2014 Power Syst. Comput. Conf.*, 2014.

[115] B. F. Hobbs, F. A. M. Rijkers, and A. F. Wals, "Strategic generation with conjectured transmission price responses in a mixed transmission pricing system — part II: application," *Power Syst. IEEE Trans.*, vol. 19, no. 2, pp. 872–879, 2004.

[116] B. F. Hobbs and F. A. M. Rijkers, "Strategic generation with conjectured transmission price responses in a mixed transmission pricing system-part I: formulation," *Power Syst. IEEE Trans.*, vol. 19, no. 2, pp. 707–717, 2004.

[117] M. van Hout, P. Koutstaal, Ö. Özdemir, and A. Seebregts, "Quantifying flexibility markets," ECN-E--14-039, Amsterdam, NL, 2014.

[118] Ö. Özdemir, P. Koutstaal, and M. van Hout, "Value of flexibility for balancing wind power generation," in *14th IAEE European Conference*, 2014, pp. 1–15.

[119] Frontier Economics, "Metis technical note T4 overview of European electricity markets," 2016.

[120] G. Morales-España, J. M. Latorre, and A. Ramos, "Tight and compact MILP formulation for the thermal unit commitment problem," *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 4897–4908, 2013.

[121] Q. Xu, S. Li, and B. F. Hobbs, "Generation and storage expansion co-optimization with consideration of unit commitment," in *Probabilistic Methods Applied to Power Systems, forthcoming*, 2018, pp. 1–6.

[122] N. Li, C. Uckun, E. M. Constantinescu, J. R. Birge, K. W. Hedman, and A. Botterud, "Flexible operation of batteries in power system scheduling with renewable energy," *IEEE Trans. Sustain. Energy*, vol. 7, no. 2, pp. 685–696, 2016.

[123] R. Sioshansi and P. Denholm, "The value of plug-in hybrid electric vehicles as grid resources," *Energy J.*, vol. 31, no. 3, pp. 1–22, 2010.

[124] R. Sioshansi, "Modeling the impacts of electricity tariffs on plug-in hybrid electric vehicle charging, costs, and emissions," *Oper. Res.*, vol. 60, no. 2, pp. 1–11, Feb. 2012.

[125] A. Tuohy, P. Meibom, E. Denny, and M. O. Malley, "Unit commitment for systems with significant wind penetration," *IEEE Trans. Power Syst.*, vol. 24, no. 2, pp. 592–601, 2009.

[126] B. M. Hodge, A. Florita, K. Orwig, and D. Lew, "A comparison of wind power and load forecasting error distributions," in *World Renewable Energy Forum*, 2012, pp. 1–8.

[127] L. Suganthi and A. A. Samuel, "Energy models for demand forecasting - a review," *Renew. Sustain. Energy Rev.*, vol. 16, no. 2, pp. 1223–1240, 2012.

[128] R. Karki, P. Hu, and R. Billinton, "A simplified wind power generation model for reliability evaluation," *IEEE Trans. Energy Convers.*, vol. 21, no. 2, pp. 533–540, 2006.

[129] Ö. Özdemir, J. de Joode, P. Koutstaal, and M. van Hout, "Generation capacity investments and high levels of renewables - the impact of a german capacity market on Northwest Europe," Energy Research Center of the Netherlands, Patten, NL, 2013.

[130] F. Wilcoxon, "Individual comparisons by ranking methods," *Biometrics Bull.*, vol. 1, no. 6, pp. 80–83, 1945.

[131] H. B. Mann and D. R. Whitney, "On a test of whether one of two random variables is stochastically larger than the other," *Ann. Math. Stat.*, vol. 18, no. 1, pp. 50–60, 1947.

[132] K. F. Weaver, V. C. Morales, S. L. Dunn, K. Godde, and P. F. Weaver, "Mann–Whitney U and Wilcoxon signed-rank," in *An Introduction to Statistical Analysis in Research: With Applications in the Biological and Life Sciences*, New York: John Wiley & Sons, Inc., 2018, pp. 297–352.

[133] "Scenario outlook and adequacy forecast 2013-2030," ENTSO-E, 2013.

[134] N. G. Singhal, N. Li, and K. W. Hedman, "A reserve response set model for systems with stochastic resources." *Working paper*.

[135] K. Van den Bergh, K. Bruninx, and E. Delarue, "Cross-border reserve markets: network constraints in cross-border reserve procurement," *Energy Policy*, vol. 113, no. November 2017, pp. 193–205, 2018.

[136] J. C. Smith, M. R. Milligan, E. A. DeMeo, and B. Parsons, "Utility wind integration and operating impact state of the art," *IEEE Trans. Power Syst.*, vol. 22, no. 3, pp. 900–908, 2007.

[137] J. M. Buchanan and W. C. Stubblebine, "Externality," *Economica*, vol. 29, no. 116, pp. 371–384, 1962.

[138] P. Lehmann, "Justifying a policy mix for pollution control: a review of economic literature," *J. Econ. Surv.*, vol. 26, no. 1, pp. 71–97, 2012.

[139] B. F. Hobbs, "What do SO2 emissions cost? Allowance prices and externality adders," *J. Energy Eng.*, vol. 120, no. 3, pp. 122–132, 1994.

[140] G. E. Metcalf and D. Weisbach, "The design of a carbon tax," *Harvard Environ. Law Rev.*, vol. 499, 2009.

[141] C. De Jonghe, E. Delarue, R. Belmans, and W. D'haeseleer, "Interactions between measures for the support of electricity from renewable energy sources and CO2 mitigation," *Energy Policy*, vol. 37, no. 11, pp. 4743–4752, 2009.

[142] S. Sorrell and J. Sijm, "Carbon trading in the policy mix," *Oxford Rev. Econ. Policy*, vol. 19, no. 3, pp. 420–437, 2012.

[143] L. Bird, C. Chapman, J. Logan, J. Sumner, and W. Short, "Evaluating renewable portfolio standards and carbon cap scenarios in the U.S. electric sector," *Energy Policy*, vol. 39, no. 5, pp. 2573–2585, 2011.

[144] P. Lehmann, "Justifying a policy mix for pollution control: a review of economic

literature," *J. Econ. Surv.*, vol. 26, no. 1, pp. 71–97, 2012.

[145] J. Bushnell, "The implementation of california AB 32 and its impact on wholesale electricity markets," *Center for the Study of Energy Markets Working Paper*, 2007. .

[146] D. Burtraw, "Regulating co2 in electricity markets: sources or consumers?," *Clim. Policy*, vol. 8, no. 6, pp. 588–606, 2008.

[147] Y. Chen, "Does a regional greenhouse gas policy make sense? a case study of carbon leakage and emissions spillover," *Energy Econ.*, vol. 31, no. 5, pp. 667–675, 2009.

[148] CAISO, "Regional integration and eim greenhouse gas compliance," *Stakeholder Processes*, 2017. [Online]. Available: www.caiso.com/informed/Pages/StakeholderProcesses/RegionalIntegrationEIMGreenhouseGasCompliance.aspx.

[149] G. Angelidis and D. Tretheway, "Eim greenhouse gas enhancement revised draft final proposal," *California Independent System Operator*. [Online]. Available: www.caiso.com/Documents/RevisedDraftFinalProposal-EnergyImbalanceMarketGreenhouseGasEnhancements.pdf.

[150] D. Tretheway, "EIM greenhouse gas enhancements 2nd revised draft final proposal," CAISO, Folsom, CA, 2018.

[151] "Program overview and design: elements of RGGI," *RGGI, Inc.*, 2018. [Online]. Available: www.rggi.org/program-overview-and-design/elements.

[152] J. Bushnell, "The implementation of california AB 32 and its impact on wholesale electricity markets," *Center for the Study of Energy Markets Working Paper*, 2007. [Online]. Available: ei.haas.berkeley.edu/research/papers/CSEM/csemwp170.pdf.

[153] L. Parker and B. D. Yacobucci, "Climate change: costs and benefits of the cap-and-trade provisions of hr 2454," Congressional Research Service 7-5700, Washington D.C., 2009.

[154] L. A. Bird, E. Holt, and G. Levenstein Carroll, "Implications of carbon cap-and-trade for US voluntary renewable energy markets," *Energy Policy*, vol. 36, no. 6, pp. 2063–2073,

2008.

[155] P. Del Río González, "The interaction between emissions trading and renewable electricity support schemes. an overview of the literature," *Mitig. Adapt. Strateg. Glob. Chang.*, vol. 12, no. 8, pp. 1363–1390, 2007.

[156] C. Gavard, N. Winchester, and S. Paltsev, "Limited trading of emissions permits as a climate cooperation mechanism? US–China and EU–China examples," *Energy Econ.*, vol. 58, pp. 95–104, 2016.

[157] C. C. Tsao, J. E. Campbell, and Y. Chen, "When renewable portfolio standards meet cap-and-trade regulations in the electricity sector: market interactions, profits implications, and policy redundancy," *Energy Policy*, vol. 39, no. 7, pp. 3966–3974, 2011.

[158] F. A. Wolak, J. Bushnell, and B. F. Hobbs, "Opinion on 'Load-based and source-based trading of carbon dioxide in California,'" Market Surveillance Committee of the California ISO, Folsom, CA, 2007.

[159] J. Kaatz and S. Anders, "The role of unspecified power in developing locally relevant greenhouse gas emission factors in California's electric sector," *Electr. J.*, vol. 29, no. 9, pp. 1–11, 2016.

[160] B. F. Hobbs, J. Bushnell, and F. A. Wolak, "Upstream vs. downstream CO2 trading: a comparison for the electricity context," *Energy Policy*, vol. 38, no. 7, pp. 3632–3643, 2010.

[161] B. F. Hobbs, J. Bushneil, and F. A. Wolak, "Is load-based carbon trading less costly than source-based trading?," *2010 7th Int. Conf. Eur. Energy Mark. EEM 2010*, 2010.

[162] J. Bushnell, Y. Chen, and M. Zaragoza-Watkins, "Downstream regulation of CO2 emissions in California's electricity sector," *Energy Policy*, vol. 64, pp. 313–323, 2014.

[163] J. Bushnell and Y. Chen, "Allocation and leakage in regional cap-and-trade markets for CO2," *Resour. Energy Econ.*, vol. 34, no. 4, pp. 647–668, 2012.

[164] M. Tanaka and Y. Chen, "Emissions trading in forward and spot markets for electricity," *Energy J.*, vol. 33, no. 2, pp. 195–221, 2012.

[165] A. Wals and F. Rijkers, "How will a CO2 price affect the playing field in the Northwest European power sector ?," in *Research Symposium European Electricity Markets*, 2003, pp. 1–11.

[166] J. Sijm, "The impact of the EU emissions trading scheme on the price of electricity in the Netherlands," Energy Research Center of the Netherlands, Amsterdam, NL, 2004.

[167] Y. Chen, J. Sijm, B. F. Hobbs, and W. Lise, "Implications of CO2 emissions trading for short-run electricity market outcomes in Northwest Europe," *J. Regul. Econ.*, vol. 34, pp. 251–281, 2008.

[168] A. Newcomer, S. A. Blumsack, J. Apt, L. B. Lave, and M. G. Morgan, "Short run effects of a price on carbon dioxide emissions from U.S. electric generators," *Environ. Sci. Technol.*, vol. 42, no. 9, pp. 3139–3144, 2008.

[169] C. Lo Prete and C. S. Norman, "Rockets and feathers in power futures markets? Evidence from the second phase of the EU ETS," *Energy Econ.*, vol. 36, pp. 312–321, 2013.

[170] A. Olson, C. K. Woo, N. Schlag, and A. Ong, "What happens in California does not always stay in California: the effect of California's cap-and-trade program on wholesale electricity prices in the Western Interconnection," *Electr. J.*, vol. 29, no. 7, pp. 18–22, 2016.

[171] C. K. Woo, A. Olson, Y. Chen, J. Moore, N. Schlag, A. Ong, and T. Ho, "Does California's CO2 price affect wholesale electricity prices in the Western U.S.A.?," *Energy Policy*, vol. 110, no. August, pp. 9–19, 2017.

[172] M. Ruth, S. A. Gabriel, K. L. Palmer, D. Burtraw, A. Paul, Y. Chen, B. F. Hobbs, D. Irani, J. Michael, K. M. Ross, R. Conklin, and J. Miller, "Economic and energy impacts from participation in the regional greenhouse gas initiative: a case study of the state of Maryland," *Energy Policy*, vol. 36, no. 6, pp. 2279–2289, 2008.

[173] California Environmental Protection Agency - Air Resource Board, "Article 5: California

cap on greenhouse gas emissions and market-based compliance mechanisms," California Code of Regulations, 2016.

[174] S. P. Holland, "Emissions taxes versus intensity standards: second-best environmental policies with incomplete regulation," *J. Environ. Econ. Manage.*, vol. 63, no. 3, pp. 375–387, 2012.

[175] M. Turner and J. Batakji, "Business practice manual for the Energy Imbalance Market, v10," CAISO, Folsom, CA, 2018.

[176] R. Mullin, "CAISO, ARB to address imbalance market carbon leakage," *RTO Insider*, 2016.

[177] CAISO, "Proposed principles for governance of a regional ISO." [Online]. Available: www.caiso.com/Documents/ProposedPrinciples-Governance-RegionalISO.pdf.

[178] CAISO, "Revised proposal principles for governance of a regional ISO." [Online]. Available: www.caiso.com/Documents/RevisedProposedPrinciples-RegionalISOGovernance.pdf.

[179] CAISO, "Regional integration California greenhouse gas compliance," *Issue Paper*, 2016. [Online]. Available: www.caiso.com/Documents/IssuePaper-RegionalIntegrationCaliforniaGreenHouseGasCompliance.pdf.

[180] E. Sauma, "The impact of transmission constraints on the emissions leakage," *Energy Policy*, vol. 51, pp. 164–171, 2012.

[181] California Air Resources Board, "California cap-and-trade program joint auction #13 summary results report," *Archived Auction Information and Results*, 2017. [Online]. Available: www.arb.ca.gov/cc/capandtrade/auction/nov-2017/summary_results_report.pdf.

[182] J. Sijm, Y. Chen, and B. F. Hobbs, "The impact of power market structure on CO2 cost pass-through to electricity prices under quantity competition - a theoretical approach," *Energy Econ.*, vol. 34, no. 4, pp. 1143–1152, 2012.

[183] W. Lise, J. Sijm, and B. F. Hobbs, "The impact of the EU ETS on prices, profits and

emissions in the power sector: simulation results with the competes EU20 model," *Environ. Resour. Econ.*, vol. 47, no. 1, pp. 23–44, 2010.

[184] K. Phillips and S. Levy, "New study shows proposed California-Pacificorp energy market integration would increase carbon pollution," *Sierra Club Press Release*, Sacramento, CA, p. 1, 20-May-2016.

[185] J. McDonald and I. Penn, "Study says expansion of California's electric grid would save consumers $1.5 billion," *Los Angeles Times*, Los Angeles, p. 1, 12-Jun-2016.

[186] "British Columbia's carbon tax," *British Columbia Climate Planning & Action*, 2018. [Online]. Available: www2.gov.bc.ca/gov/content/environment/climate-change/planning-and-action/carbon-tax.

[187] T. Sickinger, "Lawmakers unveil 'cap and invest ' carbon pricing bills," *The Oregonian*, Portland, OR, 2018.

[188] H. Bernton and P. Le, "Washington state's carbon-tax bill dies in legislature," *The Seattle Times*, Seattle, pp. 1–3, 03-Mar-2018.

[189] F. A. Wolak, J. Bushnell, and B. F. Hobbs, "Opinion on regulation energy management," Folsom, CA, 2011.

[190] J. D. Jenkins, "Political economy constraints on carbon pricing policies: what are the implications for economic efficiency, environmental efficacy, and climate policy design?," *Energy Policy*, vol. 69, pp. 467–477, 2014.

[191] "The Constitution of the United States," *Artic. 1, Sect. 8, Clause 3*.

[192] "Iberdrola renewables files at wecc to create West's 39th balancing authority," *Electricity Policy*, Online, p. 1, 16-Jun-2016.

[193] R. A. C. Van Der Veen, A. Abbasy, and R. A. Hakvoort, "A qualitative analysis of main cross-border balancing arrangements," *7th Int. Conf. Eur. Energy Mark.*, pp. 1–6, 2010.

[194] M. Burns and P. Waters, "Entergy utilities complete MISO integration," *Press Release*. [Online]. Available: www.entergy.com/news_room/newsrelease.aspx?NR_ID=2820.

[195] WECC Variable Generation Subcommittee Marketing Workgroup, "Electricity markets and variable generation integration whitepaper," 2011.

[196] R. Gramlich and M. Goggin, "The ability of current U.S. electric industry structure and transmission rules to accommodate high wind energy penetration," in *7th International Workshop on Large Scale Integration of Wind Power and on Transmission Networks for Offshore Wind Farms*, 2007, no. March, pp. 1–6.

[197] M. Milligan and B. Kirby, "The impact of balancing area size and ramping requirements on wind integration," *Wind Eng.*, vol. 32, no. 4, pp. 379–398, 2008.

[198] B. Kirby and M. Milligan, "The impact of balancing area size, obligation sharing, and energy markets on mitigating ramping requirements in systems with wind energy," *Wind Eng.*, vol. 32, no. 4, pp. 399–414, 2008.

[199] B. Kirby and M. Milligan, "Facilitating wind development : the importance of electric industry structure," Golden, CO, 2008.

[200] A. H. van der Weijde and B. F. Hobbs, "Locational-based coupling of electricity markets: benefits from coordinating unit commitment and balancing markets," *J. Regul. Econ.*, vol. 39, no. 3, pp. 223–251, Mar. 2011.

[201] P. N. Biskas, D. I. Chatzigiannis, and A. G. Bakirtzis, "European electricity market integration with mixed market designs part I: formulation," *IEEE Trans. Power Syst.*, vol. 29, no. 1. pp. 458–465, 2014.

[202] P. N. Biskas, D. I. Chatzigiannis, and A. G. Bakirtzis, "European electricity market integration with mixed market designs part II: solution algorithm and case studies," *IEEE Trans. Power Syst.*, vol. 29, no. 1. pp. 466–475, 2014.

[203] L. Meeus, L. Vandezande, S. Cole, and R. Belmans, "Market coupling and the importance of price coordination between power exchanges," *Energy*, vol. 34, no. 3, pp. 228–234, Mar. 2009.

[204] G. Oggioni and Y. Smeers, "Degrees of coordination in market coupling and counter-

trading," *Energy J.*, vol. 33, no. 3, pp. 39–90, 2012.

[205] J. Abrell and S. Rausch, "Cross-country electricity trade, renewable energy and European transmission infrastructure policy," *J. Environ. Econ. Manage.*, vol. 79, pp. 87–113, 2016.

[206] L. Vandezande, M. Saguan, L. Meeus, J. M. Glachant, and R. Belmans, "Assessment of the implementation of cross-border balancing trade between Belgium and the Netherlands," *6th Int. Conf. Eur. Energy Mark.*, pp. 1–6, 2009.

[207] S. Delikaraoglou, J. M. Morales, and P. Pinson, "Impact of inter- and intra-regional coordination in markets with a large renewable component," *IEEE Trans. Power Syst.*, pp. 1–10, 2016.

[208] B. M. Hodge, A. Florita, K. Orwig, and D. Lew, "A comparison of wind power and load forecasting error distributions," in *World Renewable Energy Forum*, 2012, pp. 1–8.

[209] J. Kazempour and B. F. Hobbs, "Value of flexible resources, virtual bidding, and self-scheduling in two-settlement electricity markets with wind generation - part I," *IEEE Trans. Power Syst.*, 2017.

[210] J. Kazempour and B. F. Hobbs, "Value of flexible resources, virtual bidding, and self-scheduling in two-settlement electricity markets with wind generation - part II," *IEEE Trans. Power Syst.*, vol. 33, no. 1, pp. 1–1, 2017.

[211] "Pcr market coupling algorithm," EPEX SPOT – Nord Pool– OMIE – OPCOM – GME – OTE – TGE, Brussels, Belgium, 2016.

[212] N. P. Padhy, "Unit commitment-a bibliographical survey," *Power Syst. IEEE Trans.*, vol. 19, no. 2, pp. 1196–1205, 2004.

[213] K. W. Hedman, "Flexible transmission in the smart grid," University of California, Berkeley, 2010.

[214] K. W. Hedman, R. P. O'Neill, and S. S. Oren, "Analyzing valid inequalities of the generation unit commitment problem," in *Power Systems Conference and Exposition, 2009. PSCE '09. IEEE/PES*, 2009, pp. 1–6.

[215] Monitoring Analytics, “2016 state of the market report for PJM - volume 1,” Independent Market Monitor for PJM, 2016.

[216] P. Etingov, Y. Makarov, R. Diao, S. Malhara, N. Zhou, R. Guttromson, J. Ma, P. Du, N. Samaan, and C. Sastry, “Analysis methodology for balancing authority cooperation in high penetration of variable generation,” Pacific Northwest National Laboratory, 2010.

[217] J. Lin and F. H. Magnago, “Pricing, modeling, and simulation of an electricity market,” in *Electricity Markets: Theories and Applications*, Piscataway, NJ: IEEE Press, 2017, pp. 211–238.

[218] R. Orans, A. Olson, and J. Moore, “WECC EDT phase 2 EIM benefits analysis & results,” Energy and Environmental Economics, Inc., prepared for Western Electricity Coordinating Council, 2011.

[219] C. Grigg, P. Wong, P. Albrecht, R. Allan, M. Bhavaraju, R. Billinton, Q. Chen, C. Fong, S. Haddad, S. Kuruganty, W. Li, R. Mukerji, D. Patton, N. Rau, D. Reppen, A. Schneider, M. Shahidehpour, and C. Singh, “The IEEE reliability test system-1996. a report prepared by the reliability test system task force of the application of probability methods subcommittee,” *Power Syst. IEEE Trans.*, vol. 14, no. 3, pp. 1010–1020, 1999.

[220] NREL, “Western wind resources dataset,” 2012. [Online]. Available: wind.nrel.gov/Web_nrel/.

[221] P. De Martini, L. Kristov, and L. Schwartz, “Distribution systems in a high distributed energy resources future: planning, market design, operation and oversight,” no. 2, p. 66, 2015.

[222] R. T. Clemen and T. Reilly, *Making hard decisions with DecisionTools*, 2nd rev. e. Pacific Grove, CA: Duxbury/Thomson Learning, 2001.

[223] B. F. Hobbs and P. Meier, “Energy decisions & the environment: a guide to the use of multicriteria methods,” in *International Series in Operations Research & Management Science*, Boston/ Dordrecht/London: Kluwer Academic Publishers, 2000, p. 257.

[224] S. A. Gabriel, A. J. Conejo, B. F. Hobbs, D. Fuller, and C. Ruiz, *Complementarity Modeling In Energy Markets*. New York: Springer, 2013.

[225] I. Staffell and M. Rustomji, "Maximising the value of electricity storage," *J. Energy Storage*, vol. 8, pp. 212–225, 2016.

[226] G. E. P. Box, "Robustness in the strategy of scientific model building," in *Army Research Office Workshop on Robustness in Statistics*, 1979, pp. 201–236.

[227] V. A. Ubhaya, "Quasi-convex optimization," *J. Math. Anal. Appl.*, vol. 116, pp. 439–449, 1986.

[228] M. Madani and M. Van Vyve, "A MIP framework for non-convex uniform price day-ahead electricity auctions," *EURO J. Comput. Optim.*, vol. 5, no. 1–2, pp. 263–284, 2017.

[229] R. Fernández-Blanco, J. M. Arroyo, and N. Alguacil, "On the solution of revenue - and network -constrained day-ahead market clearing under marginal pricing — part I : an exact bilevel programming approach," vol. 32, no. 1, p. 2551046, 2017.

[230] T. Zheng and E. Litvinov, "On ex post pricing in the real-time electricity market," *IEEE Trans. Power Syst.*, vol. 26, no. 1, pp. 153–164, 2011.

[231] W. W. Hogan, "Electricity market design and efficient pricing: applications for new england and beyond," Cambridge, MA, 2014.

[232] A. Papavasiliou and S. Oren, "Multiarea stochastic unit commitment for high wind penetration in a transmission constrained network," *Oper. Res.*, vol. 61, no. 3, pp. 578–592, May 2013.

[233] A. Papavasiliou, S. S. Oren, and R. P. O'Neill, "Reserve requirements for wind power integration: a scenario-based stochastic programming framework," *IEEE Trans. Power Syst.*, vol. 26, no. 4, pp. 2197–2206, 2011.

[234] F. Bouffard and M. Ortega-Vazquez, "The value of operational flexibility in power systems with significant wind power generation," in *Power and Energy Society General Meeting*, 2011, pp. 1–5.

[235] R. Kalaskar, "Flexible ramping product performance discussion," in *Market Surveillance Committee Meeting*, 2018, pp. 1–23.

[236] "BritNed development limited," *BritNed*, 2018. [Online]. Available: www.britned.com/.

[237] B. F. Hobbs and B. Wang, "A flexible ramping product: can it help real-time dispatch markets approach the stochastic dispatch ideal? Draft," *Electr. Power Syst. Res.*

[238] B. F. Hobbs, "Flexiramp: some economic principles needs for ramp," in *CAISO Market Surveillance Committee Meeting*, 2011.

[239] B. F. Hobbs and P. M. Meier, "Multicriteria methods for resource planning: an experimental comparison," *IEEE Trans. Power Syst.*, vol. 9, no. 4, pp. 1811–1817, 1994.

[240] D. Tretheway, "EIM greenhouse gas enhancements 2nd revised draft final proposal," CAISO, Folsom, CA, 2018.

[241] R. G. Lipsey and K. Lancaster, "The general theory of second best," *Rev. Econ. Stud.*, vol. 24, no. 1, pp. 11–32.

[242] B. H. Kim and R. Baldick, "Coarse-grained distributed optimal power flow," *IEEE Trans. Power Syst.*, vol. 12, no. 2. pp. 932–939, 1997.

[243] B. H. Kim and R. Baldick, "A comparison of distributed optimal power flow algorithms," *IEEE Trans. Power Syst.*, vol. 15, no. 2, pp. 599–604, May 2000.

[244] A. G. Bakirtzis and P. N. Biskas, "A decentralized solution to the dc-OPF of interconnected power systems," *IEEE Trans. Power Syst.*, vol. 18, no. 3, pp. 1007–1013, 2003.

[245] J. Chen, J. S. Thorp, and T. D. Mount, "Coordinated interchange scheduling and opportunity cost payment: a market proposal to seams issues," *System Sciences, 2004. Proceedings of the 37th Annual Hawaii International Conference on*. p. 10 pp., 2004.

[246] H. Seifi and M. S. Sepasian, "DC load flow," in *Electric Power System Planning: Issues, Algorithms and Solutions*, Berlin, Germany: Springer Science+Business Media, 2011, pp. 245–248.

[247] S. Hoff, "U.S. electric system is made up of interconnections and balancing authorities," *U.S. Energy Information Administration:Today in Energy*, 2016. [Online]. Available: www.eia.gov/todayinenergy/detail.php?id=27152#.

[248] E. Mengelkamp, B. Notheisen, C. Beer, D. Dauer, and C. Weinhardt, "A blockchain-based smart grid: towards sustainable local energy markets," Comput. Sci. - Res. Dev., vol. 33, no. 1–2, pp. 207–214, 2018.

VITA

Robin Sarah Broder Hytowitz was born in Los Angeles, CA. She attended Barnard College, Columbia University, graduating in 2009 with a degree in Urban Studies and Environmental Science. Simultaneously, she completed a B.A. in Women's Studies from Albert A. List College. After working for a small solar installation company, she began her master's degree at Arizona State University. She graduated in 2013 with an M.S. in Electrical Engineering, concentrating on Electric Power and Energy Systems. Robin joined the Department of Geography and Environmental Engineering at Johns Hopkins University in 2013, spending the first year of her Ph.D. in Denmark as a Fulbright Fellow. She completed her M.S. in environmental engineering in 2016 and Ph.D. in 2018. From 2015-2018 she worked as an operations research intern at the Federal Energy Regulatory Commission. Following graduation, she will join the Electric Power Research Institute as an engineer.