

Copyright
by
Yen-Yu Lee
2012

The Dissertation Committee for Yen-Yu Lee
certifies that this is the approved version of the following dissertation:

**Improving Electricity Market Efficiency: from Market
Monitoring to Reserve Allocation**

Committee:

Ross Baldick, Supervisor

W. Mack Grady

Alexis Kwasinski

David P. Morton

Diran Obadina

R. Kevin Wood

**Improving Electricity Market Efficiency: from Market
Monitoring to Reserve Allocation**

by

Yen-Yu Lee, B.S.E.; M.S.E.

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT AUSTIN

May 2012

To my family

Acknowledgments

This dissertation would not have been possible without the strong support of my advisor, committee members, colleagues, and family. It is hard to overstate my gratitude to my advisor, Dr. Ross Baldick. He gave me an opportunity to work with him even though I knew almost nothing about power systems and electricity markets at the early stage of my doctoral study. His guidance not only assisted me in completing this dissertation, but also helped my development as an engineer and researcher. I will miss the pleasant experience of working with such a brilliant and supportive advisor.

I would like to thank all my other committee members who kindly offered their help in a number of ways. Dr. David Morton gave me invaluable advice on the stochastic optimization model. I also benefited enormously from his optimization courses. Dr. Diran Obadina provided insightful comments that ensured my research was relevant to industry. Dr. Kevin Wood helped me significantly improve the presentation of this dissertation. Dr. Mack Grady and Dr. Alexis Kwansinski both offered constructive and useful advice for my research.

I am also indebted to my colleagues in the research group. Interaction and discussion with them was always rewarding. I am especially grateful to Jin Hur for his contribution to the implementation of market power indices. I

would also like to thank the Electric Reliability Council of Texas for providing financial support during my doctoral study.

I wish to thank all the friends I met during my graduate study. They always delivered timely help that I often needed. I will definitely miss the delightful times of sharing great food, watching sports, and playing softball with these friends. They made my life in graduate school much more enjoyable than I imagined it could be.

I cannot find words to express my gratitude to my family. My parents, Chong-Yin Lee and Lin-Yin Chen, have supported me in pursuing my dream with all their love. I must apologize to them for being so far away from home in the past few years. My brother, Kan-Hwa Lee, has been a role model and inspired me since I was little. My grandfather, who passed away last year, always gave me much confidence with warm encouragement. I know he will be very happy for me in heaven.

I owe my deepest gratitude to my fiancée, Orange Liu. Getting to know her is the best thing that happened during my doctoral study. I also must apologize to her for spending much more time working on my dissertation than being with her. Her patience and love helped me through the toughest moments along the way.

Improving Electricity Market Efficiency: from Market Monitoring to Reserve Allocation

Publication No. _____

Yen-Yu Lee, Ph.D.

The University of Texas at Austin, 2012

Supervisor: Ross Baldick

This dissertation proposes new methods to improve the efficiency of electricity markets with respect to market monitoring and reserve allocation. We first present new approaches to monitor the level of competition in electricity markets, a critical task for helping the markets function smoothly. The proposed approaches are based on economic principles and a faithful representation of transmission constraints. The effectiveness of the new approaches is demonstrated by examples based on medium- and large-scale electric power systems. We then propose a new system-operation model using stochastic optimization to systematically allocate reserves under uncertainty. This model aims to overcome the difficulties in both system and market operations caused by the integration of wind power, which results in a higher degree of supply uncertainty. The numerical examples suggest that the proposed model significantly lower the operation costs, especially under high levels of wind penetration.

Table of Contents

Acknowledgments	v
Abstract	vii
List of Tables	xi
List of Figures	xii
Chapter 1. Introduction	1
1.1 Overview of restructured electricity markets	2
1.2 Challenges in market monitoring	7
1.3 Challenges in reserve allocation	11
1.4 Dissertation overview	16
Chapter 2. Generator-Based Market Power Indices	18
2.1 Current Approaches to Analyzing Market Power	19
2.1.1 Principles-based analysis of market power in the absence of transmission constraints	21
2.1.2 Principled analysis of radial transmission constraints . .	27
2.1.3 Ad hoc analyses of market power with transmission con- straints	29
2.2 Transmission-Constrained Market Power Indices	31
2.3 Implementation	35
2.3.1 PowerWorld OPF	35
2.3.2 Market power indices calculation in Matlab	37
2.3.3 Visualization	38
2.3.4 Piecewise-constant offer functions	38
2.4 Case studies	41
2.4.1 IEEE 118-Bus Reliability Test System	41

2.4.2	ERCOT system	47
2.5	Conclusion	50
Chapter 3.	Firm-based Market Power Indices	54
3.1	Introduction	54
3.2	Firm-based Transmission-Constrained Market Power Index . .	56
3.2.1	Principles-based analysis of firm-based market power . .	56
3.2.1.1	Ignoring generator capacity constraints	57
3.2.1.2	Considering binding generator capacity constraints at market-clearing conditions	60
3.2.1.3	Considering binding generator capacity constraints at competitive conditions	63
3.2.2	Evaluation of market power index	66
3.2.3	Implementation	67
3.3	Case Studies of Firm-Based Transmission-Constrained Market Power Index	67
3.3.1	Single firm, no transmission constraints	67
3.3.2	Single firm in IEEE Reliability Test System	69
3.3.3	Multiple firms in IEEE Reliability Test System	73
3.3.4	ERCOT system	75
3.4	Transmission-Constrained Residual Supply Index	78
3.4.1	Definition of transmission-constrained residual supply index	79
3.4.2	Implications of transmission-constrained residual supply index	82
3.4.3	Implementation	83
3.5	Case Studies of the TCRSI	84
3.5.1	4 Bus System	84
3.5.2	IEEE 118-bus Reliability Test System	88
3.5.3	ERCOT system	89
3.6	Comparison of Indices	90
3.7	Conclusion	93

Chapter 4. Frequency-Constrained Stochastic Economic Dispatch	94
4.1 Introduction	94
4.2 Stochastic Economic Dispatch Model	98
4.2.1 Introduction	98
4.2.2 Notation	100
4.2.3 Stochastic economic dispatch model	104
4.2.4 The second-stage problem	105
4.3 Estimating Unserved Load Due to UFLS	107
4.3.1 Evaluating system frequency	108
4.3.2 Estimating the amount of UFLS	112
4.4 Decomposition Algorithm	118
4.5 Numerical Examples	121
4.5.1 Validation of simplified frequency model	121
4.5.2 IEEE 118-bus Reliability Test System	124
4.5.3 ERCOT system	127
4.5.3.1 Impact of various risk of wind ramps	129
4.5.3.2 Impact of increased level of wind integration	130
4.5.3.3 Computational costs	132
4.6 Conclusion	133
Chapter 5. Conclusion	136
5.1 Summary	136
5.2 Future Research	140
Appendices	142
Appendix A. Derivation of Subgradient of \hat{p}_ω^S	143
Appendix B. Conditions for Preserving Optimality Cuts	146
Bibliography	147
Vita	158

List of Tables

3.1	Actual and estimated mark-ups in IEEE 118-bus test system .	73
3.2	Computational performance	85
3.3	TCRSI and RSI of the four bus system	85
3.4	TCRSI and RSI of the IEEE 118-bus Reliability Test system .	88
3.5	TCRSI and RSI for selected firms in the ERCOT System . . .	90
4.1	Comparison of simulated and estimated frequency nadirs under different sizes of disturbances	122
4.2	Results for the 118 bus system	127
4.3	Estimated conditional probability of significant wind ramp event	129
4.4	Impact of various risk of wind ramps	131

List of Figures

1.1	Actual and forecasted wind power in the ERCOT system. . . .	13
2.1	Two zone network joined by radial transmission.	28
2.2	Flow chart of tool design for market power indices.	36
2.3	Piecewise-constant residual demand and its fitted quadratic curve. 39	39
2.4	Contour map of absolute value of derivative of the inverse residual demand for the 118-bus Reliability Test System.	42
2.5	Contour map of absolute value of derivative of price with respect to injection for the 118-bus Reliability Test System.	43
2.6	Contour map of estimated price-cost mark-up for the 118-bus Reliability Test System.	45
2.7	Contour map of estimated transfer of wealth for the 118-bus Reliability Test System.	46
2.8	Contour map of common log of derivative of the inverse residual demand for the ERCOT System.	48
2.9	Contour map of derivative of price respect to injection for the ERCOT System.	49
2.10	Contour map of estimated price-cost mark-up for the ERCOT system.	51
2.11	Contour map of estimated transfer of wealth for the ERCOT system.	52
3.1	Market-clearing prices for the single firm, no transmission constraint example.	69
3.2	Actual and estimated mark-ups for single firm, no transmission constraint example.	70
3.3	Contour map of firm-based TCMPI for the IEEE 118-bus Reliability Test System.	72
3.4	Contour maps of two firm-based market power analyses for the ERCOT system.	76
3.5	One-line diagram of the four bus system.	86

3.6	TCSI versus TCMI for the ERCOT system.	92
4.1	Timeline of post-contingency operation represented by the second-stage problem.	99
4.2	Illustration of constructed UFLS model.	116
4.3	Comparison of estimated and simulated frequency in time domain.	123
4.4	e_{ω}^s versus p_{ω}^s for the IEEE 118-bus Reliability Test System. . .	126
4.5	Expected cost in the ERCOT example.	132
4.6	Expected load shedding in the ERCOT example.	133
4.7	Spinning reserves procured in the ERCOT example.	134

Chapter 1

Introduction

In the past two decades, the electric power industry has been restructured around the world and more than a dozen electricity markets were created in various countries. However, there are still many problems threatening the efficiency of electricity markets. This dissertation aims to tackle two specific problems in electricity markets: market monitoring and reserve allocation. Assessing the competitiveness in electricity markets is a key procedure that helps these markets function smoothly. However, most current approaches adopted by transmission-constrained markets have only weak connections to economic principles and thus the results may be misleading. To provide more meaningful results, we propose new approaches that are based on both economic principles and faithful representations of Kirchhoff's laws. Numerical studies demonstrate that the proposed approaches indeed offer useful insights about market monitors. We then propose a new model for allocating reserve to improve both the efficiency of market operations and system reliability. Careful allocation of reserves is especially important under high levels of wind power penetration, due to the increased degree of supply uncertainty. The proposed model uses stochastic optimization to systematically allocate reserves by incorporating variability of wind, deliverability of reserves, and adequacy of

primary frequency control. Numerical results show that the proposed model can indeed improve the efficiency of system operations in terms of operation costs, compared to the current approach adopted by most electricity markets.

This chapter continues with a brief overview of restructured electricity markets. It then discusses two specific challenges addressed in this dissertation: market monitoring (Section 1.2) and reserve allocation (Section 1.3). The final section of the chapter summarizes contributions of this dissertation and provides an outline of the following chapters.

1.1 Overview of restructured electricity markets

Historically, electricity has been provided by regulated utilities that are vertically integrated. These utilities provide all required services for delivering electricity, namely, generation, transmission, and distribution. With advances in transmission technology, which enable electricity to be transported a thousand miles with less than 3% loss, the generation sector is being pulled out of the formerly integrated, three-segment industry, and no longer has the property of a natural monopoly [64]. By 1990, several regions in various countries began to deregulate the generation sector and formed wholesale electricity markets. The resulting introduction of competition is expected to make the industry more efficient and technologically innovative, thereby, reducing both costs for producing electricity and prices that consumers pay for electricity.

Electricity markets differ from other commodity markets in many aspects. For example, because of the lack of economical, large-scale energy

storage, supply and demand should balance at all times. Furthermore, the electricity market is tightly coupled with the underlying electric power grid, a large and complex engineering system. For example, the transactions of power in electricity markets must not cause any overloaded transmission line that might endanger power system reliability. Therefore, a system operator is necessary to both administer a centralized market and control the power system. In the United States, non-profit system operators are also known as independent system operators (ISOs). There are currently seven ISOs in the U.S.: California ISO (CAISO), Electric Reliability Council of Texas (ERCOT), Midwest ISO (MISO), ISO New England (ISO-NE), New York ISO (NYISO), PJM Interconnection, and Southwest Power Pool (SPP).

An ISO typically operates two centralized markets: a real-time and a day-ahead market. In a real-time market, transactions of power correspond to actual power flow in the system. The ISO clears the real-time market, that is, determines which offers for production to accept, by solving an optimal power flow model. In this case that model is referred to as “*offer-based economic dispatch*” or “*economic dispatch*.” A generator’s offer expresses the willingness-to-sell at various production quantities. According to microeconomic theory, under perfect competition, the offer would simply be the marginal cost for a generator. Economic dispatch produces price and *dispatch*, i.e., the amount of power that a ISO asks a generator to produce, for each generator while matching supply and demand with minimum total generation costs; limits on transmission line flows are imposed as constraints in economic dispatch to

ensure that no lines are overloaded.

Economic dispatch, formulated as an optimization problem, is solved by state-of-the-art optimization software [21, §4.5]. The system operator clears real-time markets every 5-15 minutes as load conditions vary. Note that economic dispatch not only produces market-clearing prices that match supply and demand, but also provides dispatch decisions for system operations. In other words, economic dispatch has the dual nature of market operations and power system operations.

The most prevalent electric power market design in North America is a locational marginal price (LMP) market, also known as a nodal pricing market. In an LMP market, the price at each location in the network reflects the marginal cost of consuming additional power at that location. An LMP market is considered the most efficient market design for managing real-time transmission congestion [64]. Also, locational prices serve as signals to incentivize generation investments at the locations where new resources are needed. This mechanism helps to keep the system efficient in the long run.

However, in an LMP market, generators which are indispensable for resolving congestion may possess *market power*, defined as the ability to profitably shift market prices from competitive levels. Market power can cause inefficient market operations or, in the worst cast, a market break-down, such as the electricity crisis in California in 2000 [33]. Therefore, reliable analyses on market power are necessary to ensure efficient market operations. The need for such analyses motivates us to develop new approaches that help *mar-*

ket monitors analyze market power, where market monitors are regulatory agencies that detect market power and take necessary actions to restore the competitiveness of the markets. We will introduce the issue of market power more thoroughly in Section 1.2, and present the new approaches for analyzing market power in Chapter 2 and 3.

Besides operations of a centralized market, the ISO is also responsible for balancing demand and supply at all times, a key issue for keeping its system stable and reliable. This task of matching demand and supply is becoming more challenging because of the increasing supply uncertainty caused by wind power generation. The need for a systematic method that helps the ISO make operational decisions under uncertainty motivates our development of a new system-operation model in Chapter 4.

In an electric power system, generators are synchronized to a single frequency, which is 60Hz in the U.S. The system operator is responsible for maintaining the system frequency within a small range to ensure system stability. This task requires a balance between generation and load at all times. To cope with load forecast errors or unanticipated loss of system components such as generators and transmission lines, the system operator needs operating reserves, that is, spare generation capacity. Operating reserves are parts of the ancillary services offered by generation providers. In some markets, such as ERCOT [21], load can also provide reserves in that certain customers can agree to reduce their power consumption, if necessary. Typical types of operating reserves are:

- **Generation regulation**, services provided by operating generators to deal with small load fluctuation. These generators adjust power output according to control signals sent from the system operator every few seconds.
- **Spinning reserves**, spare capacities provided by operating generators for unexpected loss of generators. These generators respond immediately (within several seconds) to the decline of the system frequency.
- **Non-spinning reserves**, spare capacities provided by off-line (unsynchronized) generators for replacing failed units and restoring the system to a normal state. The typical response time of non-spinning reserves is 10-30 minutes.

Operating reserves can be viewed as commodities different from energy and thus could be traded in a separate market. However, operating reserves are tightly coupled with energy, because a unit of generation capacity can provide either energy or reserve, but not both. This linkage would make separate markets inefficient, so a market should trade energy and reserves simultaneously. Under this framework, supply and demand of reserves are represented in the economic dispatch problem so that energy and reserves are co-optimized. In this formulation, the opportunity costs of selling energy or reserves are explicitly represented. Today, some ISOs operate co-optimized markets in both day-ahead and real-time markets, (e.g., MISO [40], CAISO [16], and NY-ISO [42]), while other ISOs, such as ERCOT [21], only operate co-optimized

markets in day-ahead markets.

System reliability has the property of *externality*, that is, every market participant will benefit from the increased system reliability that results from an additional unit of reserve provided by a market participant. Therefore, the demand for reserves must be regulated. ISOs are responsible for determining the appropriate allocation of reserves based on engineering analysis. Improper allocation of reserves may degrade both the efficiency in market operations and system reliability. For example, excessive reserves may result in unnecessary costs for system operations. On the other hand, insufficient reserves may endanger the system reliability and cause interruption of service for some consumers. The reserve allocation issue becomes more significant and challenging with increasing penetration levels of wind power, which causes a higher degree of supply uncertainty. The need for a systematic method to allocate reserve under uncertainty motivates our development of a new system-operation model. We will introduce the issues regarding reserves more comprehensively in Section 1.3, and present the new system-operation model in Chapter 4.

1.2 Challenges in market monitoring

Electricity markets need to be closely monitored by regulatory agencies to ensure enough competition, an essential element for well-functioning markets. *Market power*, defined as the ability to profitably shift market prices away from competitive prices [64, Section 4-1.1], can seriously weaken the levels of competition in electricity markets. Reduced levels of competition may

result in problems such as undesirable transfer of wealth from consumers to suppliers, or even lead to a market breakdown. Electricity markets typically are monitored using a variety of market power indices, which measure the competitiveness of the markets. However, many markets have adopted *ad hoc* indices that only have tenuous connections to economic incentives of exercising market power; because the foundation is *ad hoc*, the results can be misleading sometimes [64, Chapter 4-5]. This is particularly the case for the indices that consider the effects of transmission constraints, as will be explained below. Also, to better understand the strategic behavior of market participants, market monitors usually need a variety of methods that provide different insights about market power. This dissertation develops different indices to assess market power in transmission-constrained electricity markets. Unlike *ad hoc* methods, the proposed indices are based on economic principles with faithful representations of the effects of Kirchhoff’s laws.

The existence of market power results primarily from “market concentration” in the generation sector. In other words, the wholesale electricity market is typically an oligopoly, that is, this market is dominated by a limited number of large suppliers. Under this market structure, perfect competition is less likely and large suppliers usually possess some ability to control prices. In addition, the low level of price-responsive demand (i.e., only a few consumers adjust their power consumption in response to changing prices) further contributes to the market power of market participants. Market power can be exercised through either “physical” or “financial” withholding: physical with-

holding is achieved by reducing generation output, while financial withholding is achieved by raising offer prices.

The exercise of market power causes an undesirable transfer of wealth from consumers to suppliers. In the worst scenario, the excessive transfer of wealth can lead to a market breakdown, such as the electricity crisis of 2000 in California [33]. In addition, market power may lead to inefficient dispatch, which reduces total social welfare. For example, an ISO may be forced to dispatch more expensive generators when cheaper generators physically or financially withhold. To prevent the abuse of market power that leads to inefficient market operations, market power mitigation rules imposed by market monitors are necessary in an electricity market.

Market power mitigation rules should be aimed at maintaining or restoring prices to competitive levels. However, it is usually difficult to justify the appropriate levels of prices, particularly in “energy-only” markets such as ERCOT. In energy-only markets, since there is no installed capacity payment mechanism as in the restructured markets of the U.S. Northeast [51], competitive energy prices must occasionally rise above the highest typical marginal costs in the system. The occasional high prices help generation asset owners recover their investments, so that there would be enough generation capacity for growing demand. When prices are high, it is difficult to determine if these prices result from scarcity, i.e., insufficient system capacity to meet demand, or market manipulation, i.e., the exercise of market power. Failure to identify the competitive prices, i.e., the resulting prices when no market

power is exercised, may lead to inappropriate market power mitigation. On the one hand, “under-mitigation,” allowing market prices to rise excessively above competitive prices, would create an unacceptable transfer of wealth and cause sub-optimal dispatch, as discussed in the previous paragraph. On the other hand, “over-mitigation,” driving prices below competitive levels, will result in inadequate investment in generation in an energy-only market.

Appropriate market power mitigation rules should be based on market power analyses strongly connected to the fundamental economic incentives of exercising market power. Electricity markets adopted various indices as tools for quantifying market power. However, most market power indices that consider the effects of transmission constraints are only weakly connected to the economic incentives. For example, in the context of the ERCOT nodal market there is an annual competitiveness test for assessing the potential of geographical market power associated with transmission constraints [21, §3.19.1]. This test omits the economic incentives of market participants in a transmission-constrained market and thus may lead to unreliable results. On the other hand, although some indices, such as the Lerner index [64, §4-3.4] and residual supply index [59], model economic incentives, they do not incorporate the effects of transmission constraints. Such indices might also produce misleading results in LMP markets. In this dissertation, different indices are proposed to assess market power in transmission-constrained electricity markets. Unlike *ad hoc* methods, the proposed indices are built on fundamental economic principles and a more faithful representation of transmission constraints. The

numerical examples suggest that the proposed indices indeed provide useful insights about market power.

Chapter 2 describes new indices based on “small-signal” analysis, utilizing the sensitivities of prices with respect to power production quantities, under the assumption that each market participant owns all its generation assets at the same location. We relax this assumption in Chapter 3 and propose new market power indices using both “small-signal” and “large-signal” analyses: in the “large-signal” analysis, we quantify the possession of market power assuming that a generation firm withdraws all its supply from the system.

1.3 Challenges in reserve allocation

Maintaining power system reliability has been a challenging task due to the possible failure of system components and the uncertainty in demand. In restructured electricity markets, system reliability is maintained by procuring operating reserves from generation providers, with reserve requirements typically set to ensure that, for example, loss of the largest operating generator does not cause any involuntary load shedding. However, such a simple strategy for allocating reserves fails to reflect several important issues regarding reserves, such as variability of wind power, deliverability of reserves and adequacy levels of frequency control. Ignoring these issues may lead to improper allocation of reserves, which may result in unnecessary operation costs or expose the system to higher risks of interrupting electricity services. This dissertation proposes a frequency-constrained, stochastic, economic dispatch

model for allocating reserves to improve the efficiency of system operations. Unlike existing models in the literature, the proposed model addresses the key issues regarding reserves in a more comprehensive and systematic manner.

Large-scale integration of wind power poses significant challenges to both system and market operation, due to its intermittent and variable nature. Wind power is typically much more unpredictable than load, and thus significant errors in wind power forecasting are common in daily grid operation. Figure 1.1, showing the hourly-average wind power profile in the ERCOT system on April 10 of 2011, provides an example of the limitations of wind power forecasts. The solid curve (“Actual wind”) represents the actual wind power output, while the dashed (“DA wind forecast”) and dash-dotted (“HA wind forecast”) lines describe the day-ahead and hour-ahead forecasted wind power, respectively. The largest forecasting errors are around 4500MW in day-ahead forecast and 3300MW in hour-ahead forecast. In the case of an unexpected drop in wind power production, the system operator may need to deploy operating reserves to balance generation and load.

With large-scale integration of wind power, the appropriate reserve requirements are no longer obvious. According to a GE report [23], with 15GW installed wind power capacity in the ERCOT system, the size of the largest wind power ramp-down within 30 minutes would exceed 2.8GW, greater than the largest generation loss possible in the current ERCOT system (2.44GW). Also, the report suggests that the co-occurrence of both generation outages and unanticipated change in wind should be considered in the reserve allocation

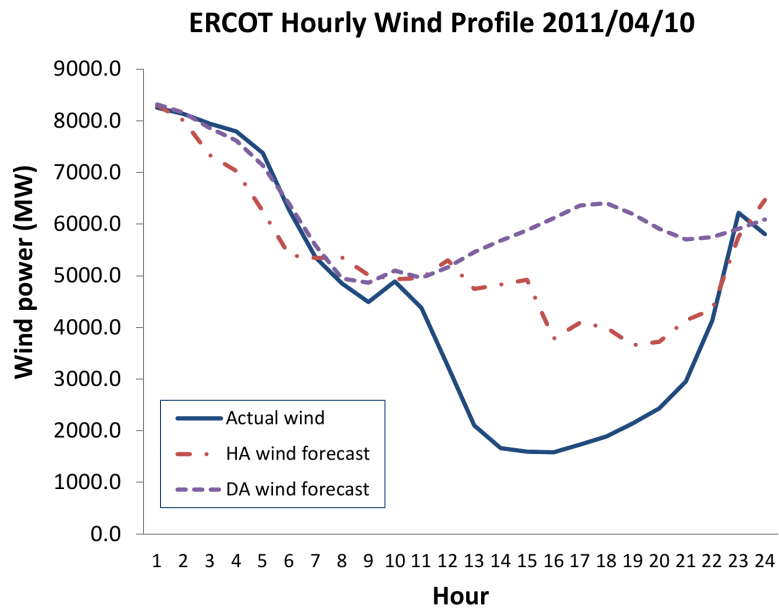


Figure 1.1: Actual and forecasted wind power in the ERCOT system. This figure indicates the limitations of wind power forecasts, exhibiting errors roughly 4500MW and 3300MW in day-ahead hour-ahead forecast, respectively. These errors are significantly greater than the capacity of the largest nuclear power plant in the ERCOT (2443MW).

process. As a result, as wind power penetration increases, the simple criterion of obtaining reserves to deal with the largest generation loss may no longer suffice.

Another problematic issue regarding reserves is their geographical allocation. Reserve requirements are usually system-wide constraints. As a result, such requirements ignore possible transmission congestion when reserves are deployed in response to contingencies. Although some systems, such as ISO-NE [31], NYISO [42], and MISO [40], use zonal reserve requirements to address this issue. However, this method may not be sufficiently accurate due to lack of representation of the meshed structure of typical transmission systems. In addition, zonal reserve requirements, usually determined in off-line studies, do not have the flexibility of being adjusted according to actual line flows in the real-time.

In addition to post-contingency transmission constraints, another important issue in reserve allocation is the sufficiency of primary frequency control, the autonomous response by generation resources to the deviation of frequency from nominal. In electric power systems, most generators are synchronized, that is, they rotate at the same frequency. This frequency will change when generation and load are not balanced: when generation output is greater (less) than load, generators speed up (slow down) causing the system frequency to increase (decrease). To maintain system reliability, it is crucial to keep the system frequency within a small range. This task is particularly challenging when the system suffers from a large disturbance, for example, an

unexpected outage of a base-load generator. A large disturbance could cause a rapid decline in system frequency, which may result in involuntary load shedding, i.e., interruption of electricity services for some customers. To deal with significant excursions in the system frequency, the system operator needs sufficient primary frequency control by having reserves allocated properly across each generating unit.

Typically, the issue of frequency control is more significant in small, isolated power systems, such as the systems in Ireland and New Zealand, because of their small system inertia [66] [67]. In such systems, the levels of inertia and reserve requirements are set to ensure that the frequency deviations are limited after the loss of the largest unit [20] [66]. With increasing levels of wind power penetration, frequency excursions might also be problematic even in large-scale isolated systems such as ERCOT, since wind resources provide only limited inertia and frequency control. Reference [36] also reports that the quality of frequency control in the U.S. has been declining in recent years and suggests that careful planning of primary frequency control is necessary.

To summarize, the main issues regarding reserve allocation include

- variability of wind power generation,
- geographical allocation of reserves, and
- sufficient primary frequency control.

Existing models [11, 35, 49, 58] for allocating reserves mostly focus on the first

two issues above and thus omit the impact of frequency control on system operation. This dissertation proposes a new economic dispatch model that systematically addresses the above issues simultaneously. This model determines an optimal energy dispatch and reserve allocation so that the costs and benefits of reliability are balanced. A more comprehensive literature surveys and the development of the model will be reported in Chapter 4.

1.4 Dissertation overview

This dissertation develops new methods to improve efficiency of electricity markets with respect to market monitoring and reserve allocation. These methods include

- new approaches, based on economic principles and Kirchhoffs' law, to analyze market power in transmission-constrained electricity markets , and
- a new economic dispatch model to dispatch generation and allocate reserves under uncertainty, in order to deal with the supply uncertainty caused, in part, by growing penetration of wind power.

The rest of this dissertation is organized as follows. Chapter 2 describes new market power indices using a “small-signal” analysis, which is based on sensitivities of prices with respect to production quantities, under the assumption that each market participant owns all its generation assets at the same location. We relax this assumption in Chapter 3 and propose new market power indices

using both “small-signal” and “large-signal” analyses: in the “large-signal” analysis, we quantify the possession of market power assuming a generation firm withdraws all its supply from the system. Chapter 4 describes the new economic dispatch model using stochastic optimization. Chapter 5 concludes this dissertation and discusses possible directions of future research.

Chapter 2

Generator-Based Market Power Indices

Electricity markets adopt various indices to assess the level of competition in electricity markets. Typically, these indices quantify market power, defined as a market participant's ability to profitably change prices from competitive levels. Market monitors, regulatory agencies that take necessary actions to maintain the competitiveness of the markets, use these indices as screening tools to identify participants that are more likely exercising their market power. However, none of these indices represent both economic incentives and the impacts of transmission constraints in principled manners. As a result, these indices may lead to inappropriate market-power mitigation that threatens the efficiency of electricity markets. To overcome the shortcomings of existing indices, we introduce four new indices of market power in transmission-constrained electricity markets that are based on economic principles and a more faithful representation of the effects of Kirchhoff's laws. The indices are developed using a "small-signal" analysis based on sensitivities of prices with respect to production quantities. We assume that each market participant owns one or more generators at the same location. This assumption will be relaxed in Chapter 3 to deal with typical case in locational marginal pricing (LMP) markets, that is, assets are usually owned at different locations.

This chapter is organized as follows. Section 2.1 surveys current approaches to analyzing market power. Then, Section 2.2 develops the indices. Section 2.3 discusses implementation of the indices. Section 2.4 presents case studies to illustrate the indices and Section 2.5 concludes.

2.1 Current Approaches to Analyzing Market Power

Current market power assessment approaches and screening tools fall into two broad categories:

1. principles-based approaches that explicitly examine incentives to deviate from competitive prices, and
2. *ad hoc* approaches based on indices such as the Herfindahl-Hirschman Index (HHI).

As an example of the first category, [64] and [70] use the derivative of the residual demand faced by a generator to assess the incentives of a hypothetical profit-maximizing firm to mark-up prices above marginal costs; however, the analysis does not explicitly treat transmission constraints. An example of this kind of assessment is outlined in section 2.1.1 in order to provide background and basis for extension to the case of transmission constraints.

As another example of the first category, [10] considers two zones joined by a single radial transmission line in a Cournot framework. While a two zone model is appropriate for some particular cases, such as modeling of California

represented as two zones (i.e., north and south), LMP markets have more complicated interactions due to the meshed nature of transmission. To summarize, the approaches for assessing market power that are rooted in economic analysis have not been extended to consider the implications of Kirchhoff's laws in a realistic meshed network. The limitations of a model of transmission that is limited to a single radial transmission line will be explored in Section 2.1.2.

Another type of principled analysis involves modeling the generation firm's behavior as a two level optimization problem [26]: the top level is the firm's profit-maximization problem and the bottom level is an optimal power flow (OPF) program that determines the market outcome, including the LMPs, given the action of the firm. While this is a principled method of modeling the strategic behavior of a firm, this two-level optimization problem is non-convex and difficult to solve, especially for large-scale systems.

As the last example of the first category, [59] and [60] discuss an index, the "residual supply index" (RSI), that reflects the degree to which a firm's offer is necessary to meet demand. At high levels of demand, physical withholding in capacity would lead to infeasibility. In this case, the generator is sometimes said to be "pivotal" [59] [60] [24] [44]. If a generator is pivotal then, in principle, it can increase the price arbitrarily. In addition, reference [44] also shows the link between the RSI and market power.

The second category, the *ad hoc* approaches, includes application to transmission-constrained markets with meshed networks. However, because the foundation is *ad hoc*, the results can be unreliable or even misleading.

(See, for example, the discussion in [64, Chapter 4-5].) An example of such an *ad hoc* screening tool, based on the ERCOT nodal market design, is outlined and evaluated critically in Section 2.1.3.

2.1.1 Principles-based analysis of market power in the absence of transmission constraints

As discussed above, [64] [70] describe an approach to assessing the incentives for a generator to mark-up its offer price above marginal costs in the absence of transmission constraints. The basic analysis considers the residual demand, that is, actual demand minus the supply of all the other participants in the market, and asks how a hypothetical profit-maximizing market participant would have offered in response to this residual demand. If that offer would involve significant mark-up of price above marginal cost and the generator is not operating at its maximum capacity, then the market participant has market power. Such a finding could be used in a subsequent market-power mitigation process.

To understand the residual demand, suppose that the demand in a particular pricing interval is D . (We ignore price-responsiveness of demand here, but it can be incorporated into the analysis.) Furthermore, consider a particular market participant k and suppose that the total offered generation of all the *other* market participants besides k is specified by the function $q_{-k} : \mathbb{R} \rightarrow \mathbb{R}$. In particular, at the price P , the total offered generation of all the other market participants is $q_{-k}(P)$. (We will follow the “economics”

convention of using the symbol P for price and the symbol q for quantity produced, in this case quantity of electricity.) The *residual demand* faced by market participant k is simply $D - q_{-k}(P)$. The inverse of the function $(D - q_{-k})$ is the inverse residual demand function faced by participant k , $p_{-k} : \mathbb{R} \rightarrow \mathbb{R}$.

Following, for example, [64] [70], to analyze the incentives faced by market participant k , we consider the conditions for market participant k to maximize its profit. In this context, profit, $\pi_k : \mathbb{R} \rightarrow \mathbb{R}$, is defined as operating profit, meaning revenue minus costs, given a generation quantity. Ignoring forward contracts, revenue equals the product of:

- quantity, q_k , of electricity produced by the generator, multiplied by
- the resulting price $p_{-k}(q_k)$,

noting that the definition of inverse residual demand is the resulting market clearing price in the market given that the generator produces the quantity q_k . We assume that the production cost function of participant k is specified by the function $c_k : \mathbb{R} \rightarrow \mathbb{R}$. Summarizing, profit for market participant k is:

$$\forall q_k \in \mathbb{R}, \pi_k(q_k) = q_k p_{-k}(q_k) - c_k(q_k).$$

Assuming that sufficient conditions for maximization are satisfied, that p_{-k} and c_k are differentiable, and that generation capacity constraints are not binding at the profit-maximizing condition, we can find the maximum of profit

by setting its derivative to zero:

$$0 = \frac{\partial \pi_k}{\partial q_k}(q_k) = p_{-k}(q_k) + q_k \frac{\partial p_{-k}}{\partial q_k}(q_k) - c'_k(q_k),$$

where $c'_k = \frac{\partial c_k}{\partial q_k}$ is the marginal cost. Re-arranging, we obtain the price-cost mark-up of price above marginal cost under the above assumptions and the hypothesis that the generator is maximizing its profits:

$$p_{-k}(q_k) - c'_k(q_k) = -q_k \frac{\partial p_{-k}}{\partial q_k}(q_k). \quad (2.1)$$

(Since the market-clearing price is non-increasing in increasing generation by market participant k , we have that $\frac{\partial p_{-k}}{\partial q_k}(q_k) \leq 0$ and the right-hand side of (2.1) is non-negative.) As a basic measure of market power, if the right-hand side of (2.1) is “large” then a profit-maximizing generator has incentives to drive up prices by withholding (at least in the absence of a forward contract, to be discussed below).

The estimate of mark-up (2.1) provides one basic index of market power that could be utilized by a market monitor. For example, a threshold could be established of, say, 10% above marginal cost or \$20/MWh above marginal cost. Any generator that is not at full production but such that the right-hand side of (2.1) is above the threshold would be flagged by market monitors as of concern of exercising market power. Such generators might then be subject to market-power mitigation such as limits on offer prices.

Equation (2.1) is sometimes re-arranged to obtain an expression for the Lerner index, an index defined as the relative price-cost mark-up with

respect to price, that is, $(p_{-k}(q_k) - c'_k(q_k))/p_{-k}(q_k)$ [64, §4-3.4]. However, as argued in [64, §4-5.2], mark-up above competitive prices (and mark-up above marginal costs) is likely to be more meaningful than the Lerner index for measuring market power. In any case, the subsequent analysis in this chapter is most easily written in terms of the price mark-up, so the rest of the analysis will focus on mark-up rather than Lerner index.

To the extent that marginal costs roughly represent the level of competitive prices, the estimated mark-up approximates the excess transfer of wealth, over and above competitive levels, from consumers to producers, per MW of production [64, §4.1]. Multiplying by production q_k , the following can be used as an index to estimate the excess wealth transfer to participant k :

$$-(q_k)^2 \frac{\partial p_{-k}}{\partial q_k}(q_k). \quad (2.2)$$

However, this index of excess wealth transfer should be used with caution since the marginal cost $c'_k(q_k)$ of participant k at its production level may be below the competitive price [64] and, consequently, the actual excess wealth transfer may be less than implied by (2.2).

So far the analysis has not considered the impacts of forward contracts, agreements between market participants to purchase or sell electricity at an agreed price and quantity on a specified time in the future. Forward contracts change the competitive situation somewhat [64, §4-4.3]. In the case that the generator has a forward contract for quantity q_k^f at price p_k^f then the profit

function becomes:

$$\forall q_k, \pi_k^f(q_k) = (q_k - q_k^f)p_{-k}(q_k) + q_k^f p_k^f - c_k(q_k).$$

Again assuming that sufficient conditions for maximization are satisfied, that p_{-k} and c_k are differentiable, and that generation capacity constraints are not binding at the profit-maximizing condition, we can find the maximum profit given a signed forward contract by setting the derivative with respect to q_k to zero:

$$0 = \frac{\partial \pi_k^f}{\partial q_k}(q_k) = p_{-k}(q_k) + (q_k - q_k^f) \frac{\partial p_{-k}}{\partial q_k}(q_k) - c'_k(q_k).$$

Again re-arranging, we obtain the price-cost mark-up with a forward contract under the hypothesis that the generator was maximizing its profits:

$$p_{-k}(q_k) - c'_k(q_k) = -(q_k - q_k^f) \frac{\partial p_{-k}}{\partial q_k}(q_k). \quad (2.3)$$

Similarly, excess wealth transfer can be estimated by:

$$-(q_k - q_k^f)^2 \frac{\partial p_{-k}}{\partial q_k}(q_k). \quad (2.4)$$

To the extent that a significant fraction of the production is forward contracted, these expressions show that the short-term incentives for mark-up are reduced [64, §4-4.3]. (However, in general, a more complicated analysis involving the interaction between incentives in the forward and “spot” markets is necessary to fully elucidate incentives [1] [43] [4] [13]. Although forward contracts do contribute to reducing incentives for exercising market power, the results are not completely understood in the context of supply offers in

an electricity market, however, and we will ignore this interaction in the discussion below by simply assuming that the expression on the right-hand side of (2.3) is appropriate as an index for assessing market power.)

Paralleling the previous argument, the right-hand side of (2.3) provides an index for assessing the incentives to exercise market power. However, it relies on knowledge of forward market positions. In some cases, forward contract positions are available to the market monitor. For example, in the context of a real-time market, the positions from the day-ahead market constitute forward financial positions.

To summarize, given the derivative of the inverse residual demand faced by a market participant, it is possible to evaluate the incentives to mark up price above marginal cost. This incentive can be used as an index of market power that, together with a threshold value of “large” for the index, constitutes a screening tool for market power. Unlike other indices such as HHI as applied to electricity markets (particularly “HHI” based on *capacity* rather than market share [64, §4-3.3]) this index is based on a principled analysis of the underlying economic incentives.

In addition to estimating the price-cost mark-up for generator k , three further indices of market power that we will show to be useful:

- the derivative of the inverse residual demand itself faced by generator k ,

$$\frac{\partial p_{-k}}{\partial q_k}(q_k),$$

- the derivative of price with respect to injection at bus k , which can be calculated from the derivative of the inverse residual demand by adding to it the slope of the offer by generator k , and
- the estimated excess wealth transfer to generator k , (2.2) or (2.4).

The analysis so far does not consider the issue of transmission constraints. No extensions of first principles analysis to the case of transmission constraints in *meshed* systems have been reported in the literature. In Section 2.2, an approach will be proposed to extending this principled analysis to a meshed transmission system.

Before discussing the proposed approach to meshed systems, however, we will first discuss the case of a single *radial* transmission constraint in Section 2.1.2. Although a single radial transmission constraint is an unrealistic situation in the context of an LMP market, it will help to explain and motivate the more complex and realistic context of a meshed system. Then, in Section 2.1.3, we will discuss existing approaches to treating market power in the presence of transmission constraints in a meshed system.

2.1.2 Principled analysis of radial transmission constraints

In a system with a single transmission line joining two zones, as shown in Figure 2.1, whenever there is transmission congestion between the zones, the two zones are separated into two markets. In this case, analysis of residual demand involves considering each zone separately. Consequently, it is straight-

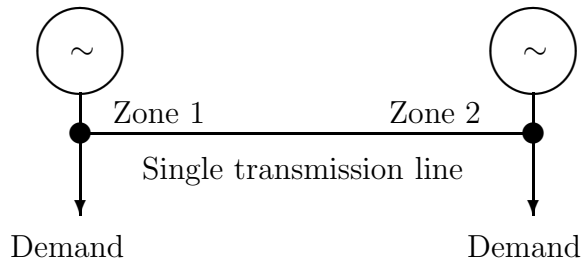


Figure 2.1: Two zone network joined by radial transmission. For a radial line, the absolute value of shift factors is always either zero or one. Consequently, given that the transmission constraint is binding, the two zone system can be validly analyzed as two separate markets.

forward to extend the approach in [64] [70] to analyze the residual demand in each zone when the single constraint is binding.

Such analysis is possible in radial systems because of a particular feature of the “shift factors,” that is, the fraction of power flowing on a line due to injection at one zone and withdrawal at another zone. For a radial line, the absolute value of shift factors is always either zero or one. Consequently, given that the transmission constraint is binding, the two zone system can be validly analyzed as two separate markets.

However, in meshed systems, the shift factors are typically between minus and plus one, so that market participants *cannot* be unequivocally partitioned into zones. The complexity of meshed systems has prompted *ad hoc* approaches that will be discussed in the next section.

2.1.3 Ad hoc analyses of market power with transmission constraints

In contrast to the discussion in Section 2.1.1, which started with an underlying economic model of profit maximization to define an index of market power, this section will consider examples of *ad hoc* analyses that attempt to analyze market power in the presence of transmission constraints in a meshed system. The lack of representation of economic incentives of these *ad hoc* methods motivates us to develop new indices that extend first principles analysis to the case of transmission constraints.

In the ERCOT nodal market there are several *ad hoc* methods are used to to assess market power. One of them, the “Element Competitiveness Index” (ECI), is styled as an *ex ante* test of competitiveness in the face of transmission constraints. In particular, it is performed to “[d]etermine if there is sufficient competition to resolve the [transmission] constraints on the import and export side” [21, §3.19]. There are several parts to this complicated test involving capacity of generators and shift factors for potentially binding transmission constraints. To summarize, the main features of the ECI test are:

- it is based on capacity of generators, and
- it is based on shift factors, and
- it considers each transmission constraint separately.

At its heart, the ECI test is an HHI test based on capacity. As Stoff points out [64, §4-3.3], despite the use of capacity-based HHI measures by

various regulatory authorities, there is no theoretical justification for capacity-based HHIs as a measure of market power. When HHIs are based on market shares instead of capacities, there is a connection to the Cournot model, where all suppliers choose, without coordination, the levels of production to maximize their profit [§4-3.2] [64]. However, even such share-based HHIs are unreliable as a measure of market power since they omit consideration of supply and demand elasticity and of forward contract positions, which are essential determinants of market power [64, §4-5.1] [9].

Although the ECI test incorporates, through the shift factors, a proxy for the geographical extent of the market, it fundamentally omits the drivers of market power: the “residual demand” faced by market participants and the forward contract positions. In particular, the ECI test is not based on any offer information [21, §3.19].

Moreover, the ECI test considers each line separately and therefore does not consider the effect of *interactions* between constraints on market power. As will be illustrated in the case studies in Section 2.4, typically, more than one constraint is binding in a large system. The focus in the ECI and other tests such as PJM’s “three pivotal supplier test” [50, Appendix J] is on particular *lines*. However, in fact, the key economic issue should be the incentives to market participants at particular *buses* due to potentially multiple interacting constraints. The ECI test is likely to obscure the locus of the fundamental economic incentives.

Another approach to monitoring market power is proposed in [41], and

is based on the sensitivity of the dispatch of generators to prices under transmission constraints. As an extension of [41], [38] uses revenue-price sensitivities to estimate price mark-ups above competitive levels. Note that the calculation of sensitivities in [41] does not represent the variation of price with generation as specified by the offers. However, representing the offers is essential to characterizing the incentives for a generator to exercise market power. As another example, [37] describes a sensitivity-based approach, which utilizes shift factors corresponding to binding transmission constraints, to find clusters of generators that can perturb market-clearing prices without affecting the dispatch. However, this analysis fails to consider the impacts of transmission constraints on the market outcome in a principled manner, because incentives faced by firms are not incorporated.

To summarize, these approaches that consider Kirchhoff’s laws in the assessment of market power have omitted fundamental economic incentives. As will be discussed in the next section, the indices we develop are based on analysis of incentives to market participants at each bus.

2.2 Transmission-Constrained Market Power Indices

In previous work, [72] describes calculation of derivatives of the inverse residual demand faced by a given generator at a single bus, assuming that offers are differentiable. To evaluate derivatives given a market-clearing result based on solution of an optimal power flow (OPF), we need the following information: shift factors for binding constraints, market-clearing quantity of

each generator, and generator offers. The calculation of derivatives is computationally efficient given the market-clearing results.

To be concrete, consider generator k located at bus k . (The case where there are multiple generators owned by different firms at a single bus is similar. We discuss the case where a firm owns multiple generators at different buses in Section 3.2.) The calculation of derivative of the inverse residual demand at bus k considers the effects of binding transmission constraints and quantitatively evaluates the decrease in residual demand elasticity faced by generator k when transmission constraints are binding.

Suppose that $H \in \mathbb{R}^{b \times w}$ is the matrix of shift factors for injection at marginal generators $j \neq k$ to the b binding constraints in the system at the market-clearing conditions, with bus k chosen to be the (price) reference bus and assuming that there are w marginal generators $j \neq k$. That is, $\forall \ell, H_{\ell j}$ is the amount of power flowing on constrained line ℓ given a unit injection at bus j and withdrawal at bus k . Note that the generators $g \neq k$ at their full capacities are excluded from the calculation, since their injections are assumed to be fixed in the sensitivity analysis. Furthermore, let $\Lambda \in \mathbb{R}^{w \times w}$ be a diagonal matrix whose diagonal entries are the inverses of the derivatives of the offers at buses other than bus k , evaluated at the market-clearing conditions. Then, from [72], the derivative of inverse residual demand faced by generator k evaluated at the market-clearing injection q_k^* is:

$$\frac{\partial p_{-k}}{\partial q_k}(q_k^*) = \left[-\mathbf{1}^\top \Lambda \mathbf{1} + \mathbf{1}^\top \Lambda H^\top (H \Lambda H^\top)^{-1} H \Lambda \mathbf{1} \right]^{-1}, \quad (2.5)$$

where $\mathbf{1} \in \mathbb{R}^w$ is the vector of all ones and superscript \top means transpose. That is, to calculate the derivative of the inverse residual demand for generator k , we need the following:

- The derivative of the offer curve of each generator $j \neq k$ evaluated at the market-clearing quantity, and
- The shift factor to the binding transmission lines for injection at each generator $j \neq k$ and withdrawal at generator k .

The derivative of the inverse residual demand provides the first of four proposed indices for transmission-constrained market power, and shows the level of price responsiveness faced by the generator at bus k due to the combination of offers and transmission constraints in the rest of the system. A second index is the derivative of price with respect to injection. This second index shows the level of price responsiveness at bus k including all offers and transmission constraints in the system, including the offer at bus k . It shows the level of price responsiveness faced by, for example, a new entrant at bus k .

We develop two further indices of the market conditions faced by the generator at bus k that depend on the transmission-constrained derivative of the residual demand. These indices are conceptually straightforward extensions of the development in Section 2.1.1 to the transmission-constrained case. That is, they are based on analyzing the incentives to a profit-maximizing market participant, assuming that each market participant owns generation at only a single bus.

As a third index, substitute into the right-hand side of (2.3):

- the market-clearing injection at bus k , $q_k = q_k^*$;
- the forward contract quantity, q_k^f ; and
- the derivative of the inverse residual demand evaluated at the market-clearing injection, $\frac{\partial p_{-k}}{\partial q_k}(q_k^*)$.

The expression that is derived provides the estimated price-cost mark-up, again assuming that sufficient conditions for profit maximization are satisfied, that p_{-k} and the cost c_k are differentiable, and that generator capacity constraints are not binding at the profit maximum. As a fourth index of transmission-constrained market power, excess wealth transfer can be estimated using (2.4).

In contrast to, say, ECI, the index (2.3) has a concrete interpretation in terms of market power: it estimates the mark-up of price above marginal cost for a hypothetical profit-maximizing generator. If forward contract information was not available then $-q_k \frac{\partial p_{-k}}{\partial q_k}(q_k)$ could be used as an index instead; however, any subsequent market-power mitigation should be sensitive to the implications of forward contracting on market power. The index (2.4) also has a concrete interpretation: it estimates the excess transfer above competitive levels.

To summarize, four indices of transmission-constrained market power developed in this section are:

- the derivative of the inverse residual demand (DIRD),
- the derivative of price with respect to injection (DPI),
- the estimated price-cost mark-up (EPM), and
- the estimated wealth transfer over competitive levels (EWT).

Calculation and visualization of these indices is described in the next section.

2.3 Implementation

We have implemented the prototype calculation of the four indices developed in Section 2.2 using Matlab [39] and PowerWorld [53]. Figure 2.2 shows the flow chart of the tool design. Given generation offer cost data, load data, and transmission network data, PowerWorld first solves the OPF model. With the results from the OPF solution, the market power index is evaluated using Matlab. We ignore forward contracts. The calculated market power index is then displayed on contour maps using PowerWorld to help market monitor understand the levels of competition on a geographical map. The following sections discuss implementation in detail.

2.3.1 PowerWorld OPF

Using linear programming (LP), PowerWorld solves the OPF in a DC model, which is a linear approximation of the “full” non-linear AC network model. The PowerWorld OPF implements various representations of cost

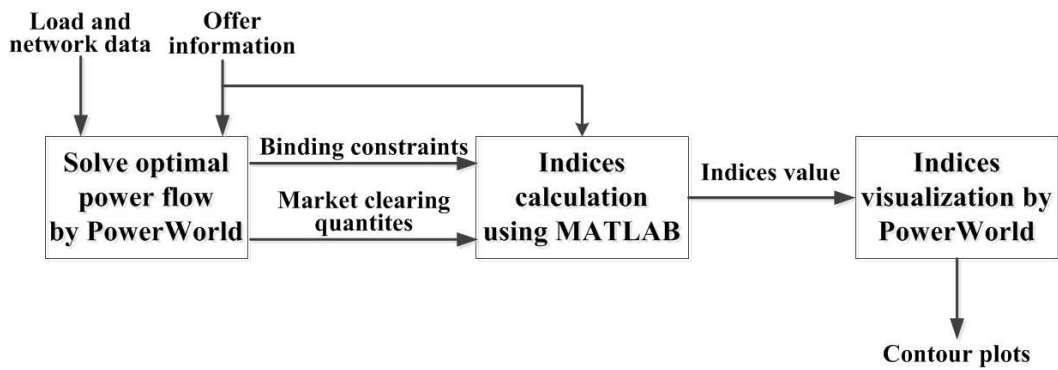


Figure 2.2: Flow chart of tool design for market power indices. Given generation offer cost data, load data, and transmission network data, PowerWorld solves the OPF model. With the results from the OPF solution, the market power index is evaluated using Matlab. The calculated market power index is then displayed on contour maps using PowerWorld to help market monitor understand the competitive situation geographically.

curves. In the context of an electricity market, the *cost* data used by the OPF corresponds to the integral of the offers. We utilize the representation in PowerWorld that allows for cubic (or lower degree) cost curves (corresponding to quadratic or lower degree offer curves). This guarantees that the offers are differentiable, consistent with the assumption in the analysis of residual demand as developed in [72]. However, the cubic cost curves require that piecewise-constant and other non-smooth offers be smoothed.¹

Generator data including cost data, transmission network data, and load data are utilized by PowerWorld to calculate the OPF. From the OPF solution, several other files are produced in order to calculate the market power index as described in more detail in the next section.

2.3.2 Market power indices calculation in Matlab

The files produced by PowerWorld are passed to Matlab for calculation of the indices. In order to avoid matrix inversion and to avoid ill-conditioning issues, *QR* factorization and forwards and backwards substitution is used to evaluate (2.5). The four indices, DIRD, DPI, EPM, and EWT, are evaluated for each generator.

¹Since LP is used in PowerWorld, the polynomial cost curves are approximated by piecewise-linear curves for calculations within PowerWorld.

2.3.3 Visualization

Market monitors are likely to benefit from figures that can show the varying levels of competitiveness at different buses on a geographical map, especially for a large interconnected system. PowerWorld contour function is a visualization tool commonly used in the industry [47]. We therefore use PowerWorld to display the indices with color contours. As shown in Figure 2.2, the computed market power indices are passed to PowerWorld for display.

2.3.4 Piecewise-constant offer functions

Electricity markets typically adopt non-differentiable offer curves, for example, ERCOT, Midwest ISO, and the Southwest Power Pool, use piecewise linear offers [21, §4.4.9.3 and 6.4.3] [40, §4.2.2.2.1] [63, §5.4]. In some markets, offers can only be specified by piecewise-constant functions. Non-differentiable offer curves, especially piecewise-constant offers, pose difficulties for calculating the indices since the offer slopes at the market-clearing quantities might be infinite or 0. As discussed in [69], due to the large number of price increments allowed for each generator, plus the large number of generators in a market, the number of steps in the residual demand curve faced by any market participant is typically large, especially in the neighborhood of the market clearing quantity. Therefore, this issue can be dealt with by fitting a smooth curve to the piecewise-constant offer function in the neighborhood of the market-clearing quantity, so that the average slope in the vicinity of the market clearing quantity is estimated.

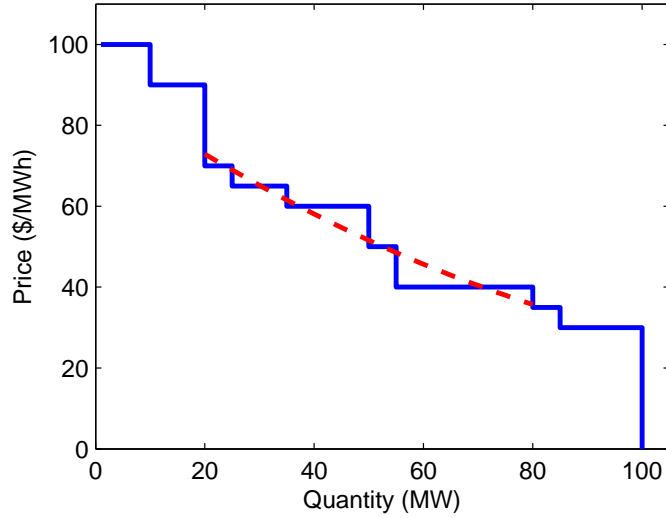


Figure 2.3: Piecewise-constant residual demand and its fitted quadratic curve near 50MW.

Although smoothing out the piecewise-constant offers introduces some errors, this approximation seems to work well given that the number of steps in the residual demand curve is large. The following example shows how smoothing out the offer does not necessarily introduce serious errors into the estimation of the mark-up. Consider generator k facing a residual demand curve as shown in Figure 2.3: \$100/MWh for $[0, 10)$ MW, \$90/MWh for $[10, 20)$ MW, \$70/MWh for $[20, 25)$ MW, \$65/MWh for $[25, 35)$ MW, \$60/MWh for $[35, 50)$ MW, \$50/MWh for $[50, 55)$ MW, \$40/MWh for $[55, 80)$ MW, \$35/MWh for $[80, 85)$ MW, and \$30/MWh for $[85, 100)$ MW. Assume that the operating cost of generator k is $0.1Q^2 + 20Q$.

We claim that the profit-maximizing quantity of the generator would

be (just less than) 50 MW, with a clearing price of approximately \$60/MWh.

To see this note that:

- if generator k were to produce more than 50 MW then the price would drop by \$10/MWh from \$60/MWh to \$50/MWh and profit would drop by around $\$10/\text{MWh} \times 50 \text{ MW}$,
- if generator k produces somewhat less than 50 MW then the price would stay the same, and the price would be above marginal cost, but the generator would sell less, lowering its profit, and
- if generator k produces significantly less than 50 MW in order to drive prices up to \$65/MWh then the production would be reduced by around 15 MW and profit would decrease by around \$300/h or more.

Given the profit-maximizing condition for generator k , its actual mark-up is \$30/MWh.

By fitting a quadratic curve to the offer in the neighborhood of 50MW (from 20MW to 80MW, see dashed curve in Figure 2.3) and estimating the slope of the fitted curve at the market-clearing quantity 50MW, the approximated slope of the offer is $-0.6196(\$/\text{MWh})/\text{MW}$. Based on this slope, the estimated price-cost mark-up is \$30.98/MWh, which is very close to the actual mark-up. This example shows that smoothing offer can give a reasonable approximation to the actual mark-up in this case.

One reason why the approximation of the derivative does not necessarily produce a large error is related to the assumption of profit maximization; the change in price and change in quantity to adjacent segments must be such as to involve a decrease in profits. This effect is well approximated by the average slope over several adjacent segments.

In our market power indices tool, non-decreasing quadratic curves are used to fit the piecewise-constant offers near the market-clearing quantities and the average offer slopes over several adjacent segments are evaluated based on the quadratic curves. Note that this post-processing step for the offer functions is only required for calculating the residual demand derivatives, while the OPF solver still uses piecewise-constant offers to clear the market.

2.4 Case studies

This section demonstrates the use of indices in the context of the IEEE 118-bus Reliability Test System and the ERCOT system. We will show that these four indices provide different insights about market power, which should assist market monitors to better understand the competitive situation in the markets.

2.4.1 IEEE 118-Bus Reliability Test System

Power flow data for the IEEE 118-bus Reliability Test System [54] together with generic cost data were used to create a data set for PowerWorld. An OPF was solved for this system with all pre-contingency thermal limits

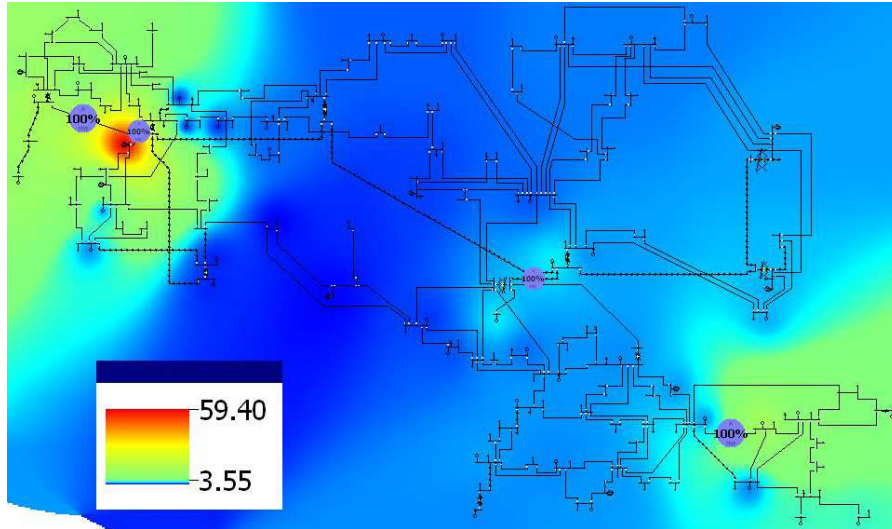


Figure 2.4: Contour map of absolute value of index DIRD for the 118-bus Reliability Test System. The units are $(\$/\text{MWh})/\text{GW}$. Many buses in the system are such that the residual demand faced at that bus is not significantly affected by the transmission constraints (in blue contours). The red contours in the upper left suggests that changes in injection at that bus would significantly affect prices.

imposed. An evaluated index for each bus is displayed in color contours on a geographical map according to a predefined mapping between index value and color. For the regions without any bus, since no index value is specified, PowerWorld just automatically interpolates the color contours in these regions [52]

Figure 2.4 shows the the contours of the index DIRD for this system, given offers equal to generic marginal costs. Note that four transmission constraints (shown by the four light blue “pie charts” labelled as 100% flow) are binding in this system.

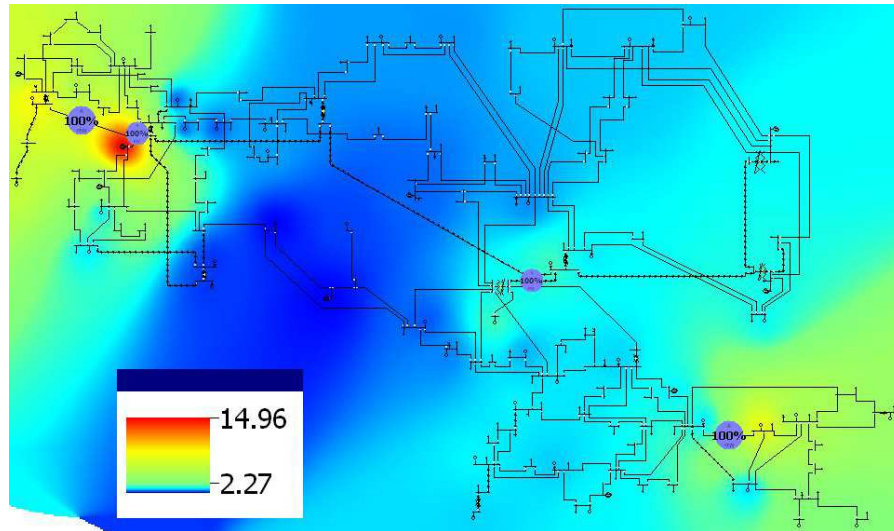


Figure 2.5: Contour map of absolute value of the index DPI on for the 118-bus Reliability Test System. The units are $(\$/\text{MWh})/\text{GW}$. The contours are qualitatively similar to those in Figure 2.4. However, the magnitudes of the derivative are smaller, reflecting the the effect of inclusion of the offers at each bus in the index DPI. This figure suggests the levels of competition faced by a new market entrant.

Figure 2.4 indicates the competitive situation faced by generators at the buses, as determined by the shift factors to binding constraints, market-clearing quantities, and the offers. The large blue regions in the figure imply that buses in these regions have the same or very close values of the DIRD. Other buses are somewhat affected by the transmission constraints, as shown by the green contours.

However, one bus in this system has a significantly greater magnitude for the index DIRD, as shown by the red contours in the upper left of the system near to two of the binding transmission constraints. This indicates that changes in injection at this bus would significantly affect prices. Generation at this bus would potentially be flagged by market monitors for further market power analysis.

Figure 2.5 shows the contours of the index DPI for the Reliability Test System. The contours are qualitatively similar to those in Figure 2.4. However, the magnitudes of the derivative are smaller, reflecting the effect of inclusion of the offers at each bus in the derivative of price with respect to injection. Figure 2.4 suggests the levels of competition that would be faced by a new market entrant.

Figure 2.6 shows the index EPM according to (2.1). Many buses have a very low estimate, as shown by the dark to light blue contours. However, it is interesting to note that even in the regions of low magnitude of DIRD, the EPM can still be moderately high, as shown by the green contours. Unsurprisingly, the bus that has a high value of magnitude for its DIRD also has a high EPM,

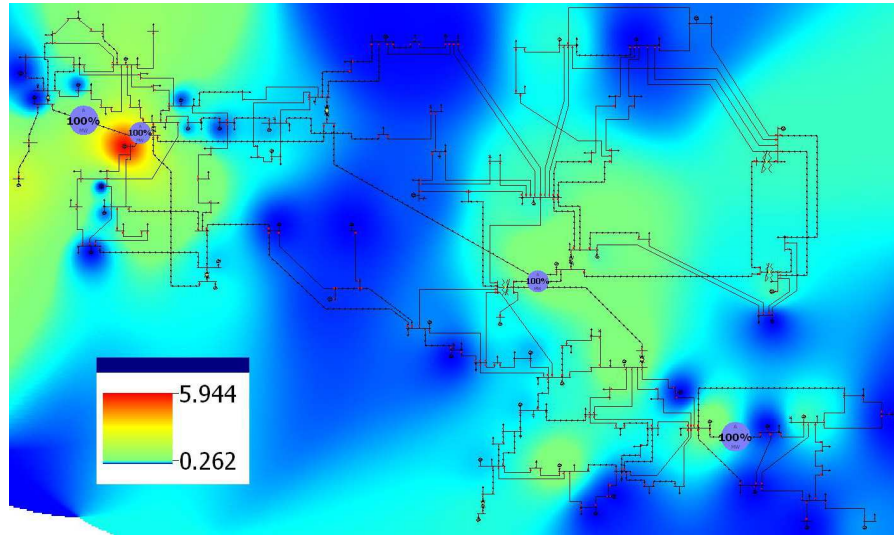


Figure 2.6: Contour map of the index EPM for the 118-bus Reliability Test System. The units are \$/MWh. Many buses have a very low estimate, as shown by the dark to light blue contours. However, it is interesting to note that even in the regions of low magnitude of DIRD (see Figure 2.4), the EPM can still be moderately high, as shown by the green contours. In other words, the EPM provide some insights about market power that are complementary to the DIRD.

as indicated by the red contours. Figure 2.7 shows the contours of the EWT according to (2.2). There is a generator (in red contours closed to the left margin of the figure) that has large EWT, but only with moderate EPM. This results imply that market monitors may need indices with different insights about market power to better understand the strategic behavior of generators.

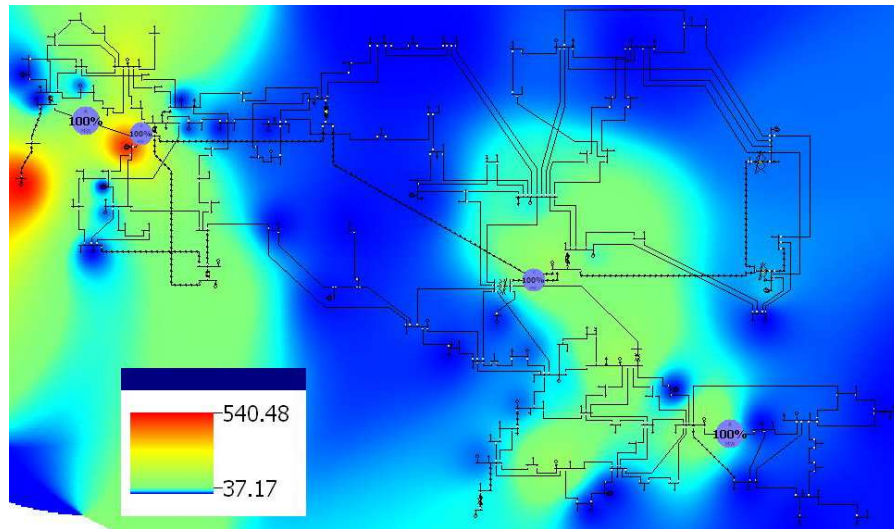


Figure 2.7: Contour map of index EWT for the 118-bus Reliability Test System. The units are \$/h. There is a generator (in red contours closed to the left margin of the figure) that has large EWT, but only with moderate EPM. This observation suggests that market monitors may need various indices to better understand the strategic behavior of generators.

2.4.2 ERCOT system

Power flow data for the ERCOT system based on summer peak in 2007, together with generic cost data² is used to create a data set for PowerWorld. A security-constrained OPF is solved for this system with all thermal pre- and single post-contingency constraints enforced except for those post-contingency constraints that could not be satisfied, such as those involving

- radially connected generators,
- radially connected loads, and
- nearby normally open circuit breakers that could be closed in the event of the contingency to mitigate an overload.

We do not enforce voltage constraints since the coupling between voltage and the market outcomes is typically weak [34].

There are three binding pre-contingency constraints and sixteen binding post-contingency constraints at the OPF solution, indicating the importance of considering multiple interacting constraints. The index DIRD is calculated from the solution of the security-constrained OPF according to (2.5) and is shown in Figure 2.8 (on a common log scale). Most areas in ERCOT have relatively small magnitudes for DIRD. However, generators in several regions,

²For instance, for a 600MW coal-fired generation unit, the marginal cost curve, represented as a linear function, ranges from \$24-30/MWh. As another example, for a 400MW gas-fired steam turbine, the marginal cost curve, also represented as a linear function, ranges from \$63-71/MWh.

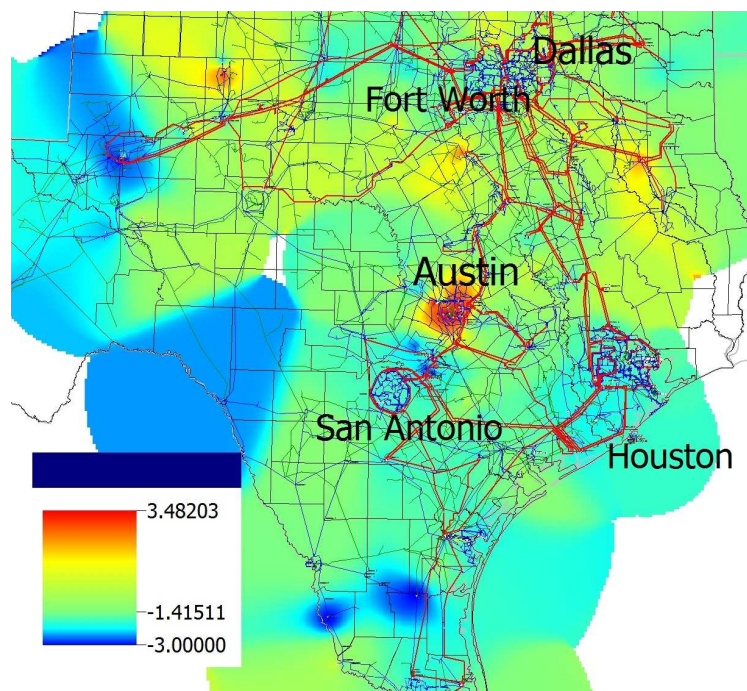


Figure 2.8: Contour map of common log of index DIRD for the ERCOT System. The units are $(\$/\text{MWh})/\text{MW}$. Most areas in the ERCOT have relatively small magnitude of DIRD. However, generators in several regions, indicated by the red contours, face quite inelastic residual demand. The red contours near Austin are particularly noticeable, presumably because several generators in the region have large shift factors to a binding transmission constraint.

indicated by the red contours, face quite inelastic residual demand. The red contours near Austin are particularly noticeable, presumably because several generators in the region have large shift factors to a binding transmission constraint. Compared with the DIRD, the index DPI has much smaller magnitude, as shown in Figure 2.9. This figure suggests that the level of competition that would be faced by a new market entrant is likely to be strong at all buses.

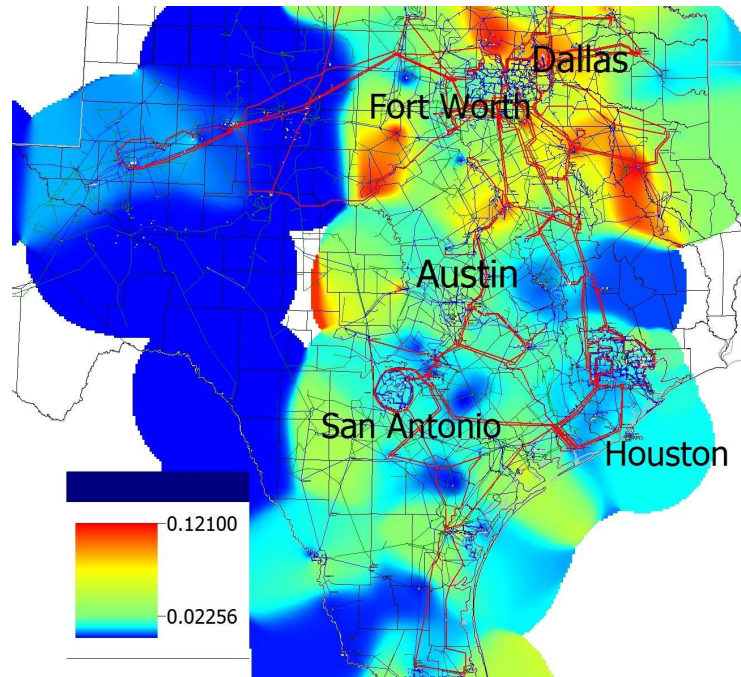


Figure 2.9: Contour map of derivative of price respect to injection for the ERCOT System. The units are $(\$/MWh)/MW$. Compared with the DIRD, the index DPI has much smaller magnitude, as shown in Figure 2.9. This figure suggests that the level of competition faced by a new entrant is likely to be strong in each bus.

Figure 2.10 shows the index EPM (on a common log scale) according to (2.1) for buses in the ERCOT system. Most of ERCOT has relatively low values for the EPM. However there are several regions of moderate EPM and one region with very large EPM near Austin. This is presumably due to binding transmission constraints impeding the import of power into these regions. The index EPM of more than 5 million \$/MWh are unrealistic given the (unmodeled) levels of forward contract cover and offer caps. If this information were known, more realistic estimates could be obtained that would more likely reveal the incentives for generators to drive up the prices. Figure 2.11 shows the index EWT (on a common log scale) according to (2.2). Some regions near Austin have moderately high EWT (in yellow contours) but small EPM (in light blue contours). This observation implies that these indices provide different insights about market power that should help market monitor better understand the competitive situation.

2.5 Conclusion

This chapter developed four indices of market power in the context of transmission-constrained electricity markets. Unlike *ad hoc* methods adopted in most electricity markets, the proposed indices incorporate economic principles and faithfully represent the effects of Kirchhoff's laws. Techniques for efficient implementation are also developed. Case studies show that these four indices provide different insights about market power and help market monitors better understand the strategic behaviors of market participants. These

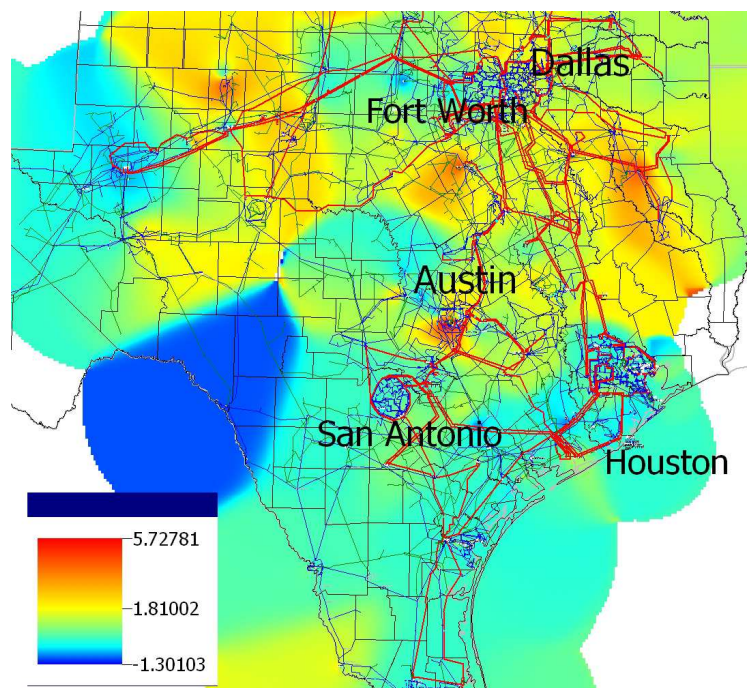


Figure 2.10: Contour map of common log of the index EPM (in \$/MWh) for the ERCOT system. Most of ERCOT has relatively low values for the EPM. However there are several regions of moderate EPM and one region with very large EPM near Austin. This is presumably due to binding transmission constraints near Austin impeding the import of power into these regions.

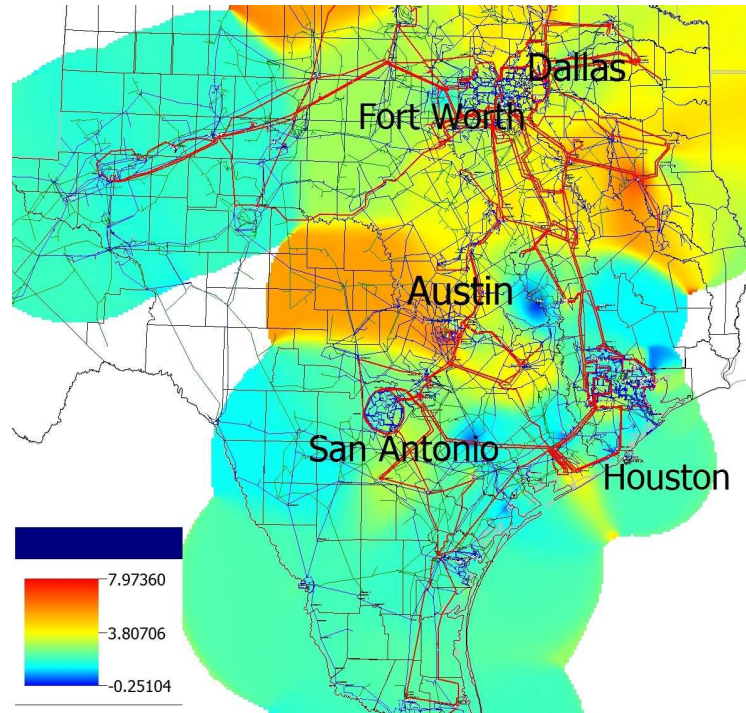


Figure 2.11: Contour map of common log of the index EWT (in \$/h) for the ERCOT system. Some regions nearby Austin have moderately high EWT (in yellow contours) but small EPM (in light blue contours). This observation implies that these indices provide different insights about market power that may help market monitor better understand the competitive situation.

indices could be evaluated in different contexts, such as:

- part of an *ex ante* simulation of market operation over pricing intervals in a time horizon using a production cost simulator, either based on competitive offers or based on some assumed strategic offers; this simulation helps market monitors predict the existence of market power; and
- part of an *ex post* analysis by a market monitor, based on historical offer information. This analysis helps market monitor apply appropriate market-power mitigation rules to restore prices to competitive levels.

The next chapter generalizes the analysis based on economic principles to a situation in which a market participant owns generation assets in different buses, which is typically the case in real-world electricity markets.

Chapter 3

Firm-based Market Power Indices

3.1 Introduction

Chapter 2 presented four new market power indices in transmission-constrained electricity markets. But one key assumption in developing those indices is rarely valid: all generators that a market participant owns are collocated. In this chapter, this assumption is relaxed and two different approaches for analyzing firm-based market power considering transmission constraints are proposed. The first one is based on “small-signal” analysis, while the other is based on “large-signal” analysis. They provide different insights into market power and can both be integrated into market power analysis. Both approaches analyze market power for a single market-clearing interval but, because of their computational efficiency, could also be applied repeatedly over multiple market-clearing intervals to assess the average.

The first approach generalizes the principled analysis proposed in Chapter 2, which is based on the the transmission-constrained residual demand derivative that was described in [72]. Unlike *ad hoc* approaches (see Section 2.1.3), the four market power indices in Chapter 2 represent the incentives faced by a firm, including the effects of Kirchhoff’s laws. That analysis is an

ex post assessment based on offer information and market-clearing results, and is computationally efficient.

A limitation of the indices in Chapter 2, however, is that each generation firm is assumed to own generation assets at only a single bus. In reality, firms often own assets located at multiple buses having different locational marginal prices (LMP). As shown by an example in [17], strategic behavior of firms with assets at multiple buses can be qualitatively different to the simpler case analyzed in Chapter 2. In this chapter, based on the Jacobian of the inverse residual demand function (JIRD) described in [73], we extend the transmission-constrained market power indices to the case that a firm has assets at multiple locations, and propose an analogous firm-based transmission-constrained market power index (TCMPI).

We also propose a second approach to assess firm-based market power: the transmission-constrained residual supply index (TCRSI), which generalizes the residual supply index (RSI) [59] [60] to the case of binding transmission constraints. The RSI was developed by the California ISO (CAISO) and is used to predict market power in the CAISO and some other markets. The RSI reflects the degree to which a firm's offers are necessary to meet demand. Empirical results indicate that the RSI is strongly correlated to the exercise of market power. However, the original RSI definition ignores the effects of transmission constraints and might be less helpful in the context of LMP markets.

The CAISO uses "competitive path assessment" to determine the com-

petitiveness of each transmission constraint [14]. However, this approach focuses on the degree of violation of transmission capacity constraints needed to meet demand without a given firm’s supply, whereas our approach directly generalizes the RSI to the case of transmission constraints. The TCRSI proposed here can efficiently assess the extent to which a firm is “pivotal” when transmission constraints bind.

The rest of this chapter is organized as follows: Section 3.2 introduces the TCMPI. Section 3.3 provides case studies of the TCMPI. Section 3.4 develops the TCRSI. Section 3.5 provides case studies of the TCRSI. Section 3.6 compares these two approaches and Section 3.7 concludes. Our discussion will not treat forward contracts explicitly, but the analysis can easily be modified to include the effect of exogenously specified forward contracts.

3.2 Firm-based Transmission-Constrained Market Power Index

To develop the firm-based transmission-constrained market power index, we first analyze the incentives for a firm to profitably alter the prices from competitive levels. Then we discuss how to compute the index efficiently and provide some examples.

3.2.1 Principles-based analysis of firm-based market power

Suppose that firm i owns generators at buses $k \in K_i$, where $|K_i| = r$. We collect the production quantities $q_k, k \in K_i$, at all these generators into

a vector $q \in \mathbb{R}^r$. Assume that $p_{-k} : \mathbb{R}^r \rightarrow \mathbb{R}$ is the resulting market-clearing price at bus k given that the firm produces the quantities q . That is, p_{-k} is the inverse residual demand function for firm i at bus k . We collect the inverse residual demands $p_{-k}, k \in K_i$, together into a vector function $p : \mathbb{R}^r \rightarrow \mathbb{R}^r$.

Note that the inverse residual demand faced by a generator at bus k depends on the whole vector q . That is, actions by the firm at any one of its generators may result in a change in the price at bus k .

We now assume that firms are profit maximizers and consider the condition for firm i to maximize its profit. Assume that the production cost functions of the firm are specified by the cost functions $c_k : \mathbb{R} \rightarrow \mathbb{R}, k \in K_i$. Ignoring forward contracts, the profit for the market participant is:

$$\forall q \in \mathbb{R}^r, \pi(q) = \sum_{k \in K_i} \left(q_k p_{-k}(q) - c_k(q_k) \right),$$

where the term $\sum_{k \in K_i} q_k p_{-k}(q)$ is the total revenue of firm i . (The case with exogenously specified forward contracts is similar.) We first consider the case where the capacity constraints of the generators owned by firm i are not binding and then consider the more general case.

3.2.1.1 Ignoring generator capacity constraints

Assuming that sufficient conditions for maximization are satisfied, that p and $c_k, k \in K_i$, are differentiable, and that generator capacity constraints of firm i are not binding, we can maximize the profit of firm i by setting the partial derivatives of profit with respect to quantities equal to zero. Focusing

on the partial derivative with respect to q_m for a particular $m \in K_i$, we obtain:

$$0 = \frac{\partial \pi}{\partial q_m}(q) = p_{-m}(q) + \sum_{k \in K_i} q_k \frac{\partial p_{-k}}{\partial q_m}(q) - c'_m(q_m),$$

where $c'_m = \frac{\partial c_m}{\partial q_m}$ is the marginal cost of the generator owned by firm i at bus m . Re-arranging the above equation, we obtain the price-cost mark-up at bus m under the hypothesis that the firm was maximizing its profits:

$$p_{-m}(q) - c'_m(q_m) = - \sum_{k \in K_i} q_k \frac{\partial p_{-k}}{\partial q_m}(q). \quad (3.1)$$

This is a generalization of (2.1) to the case of firms owning multiple generators. In the firm-based analysis, the profit-maximizing mark-up at generator m depends on productions at other buses q_k and cross derivatives $\frac{\partial p_{-k}}{\partial q_m}$.

To aggregate all the estimated mark-ups into a single index for firm i , it is natural to evaluate the quantity-weighted average mark-up of the firm:

$$\frac{\left(- \sum_{m \in K_i} q_m \sum_{k \in K_i} q_k \frac{\partial p_{-k}}{\partial q_m}(q) \right)}{(\mathbf{1}^\top q)} = \frac{\left(-q^\top \frac{\partial p}{\partial q}(q) q \right)}{(\mathbf{1}^\top q)}, \quad (3.2)$$

where: $\frac{\partial p}{\partial q}(q)$ is JIRD evaluated at q , superscript \top means transpose, and $\mathbf{1} \in \mathbb{R}^r$ is the vector of all ones. Note that the matrix $\frac{\partial p}{\partial q}(q)$ is symmetric and negative semi-definite, as proved in [6] and [73]. Therefore, the estimated average mark-up is always non-negative.

Since cross-derivatives $\frac{\partial p_{-k}}{\partial q_m}$ for $k \neq m$ can be positive, it may be the case that, at some buses, profit maximization corresponds to a *mark-down*

rather than a mark-up. That is, for some buses m , the estimated mark-up in (3.1) may be negative. While this seems to be counter-intuitive, Hogan [27] and Cardell *et al.* [17] describe just such a case where a firm offers below marginal cost at bus e on the exporting side of a constraint in order to congest the line and consequently be able to offer well above marginal cost at a bus m on the importing side. That is, $-\sum_{k \in K_i} q_k \frac{\partial p_{-k}}{\partial q_e}(q)$ is negative, while $-\sum_{k \in K_i} q_k \frac{\partial p_{-k}}{\partial q_m}(q)$ is significantly positive. The mark-up at each bus considered separately does not, in this case, give a full picture of the situation faced by a firm. However, the index (3.2) combines the effect of mark-up at all buses.

To summarize, the TCMPI proposed here is the quantity-weighted average of estimated price-cost mark-ups of firm i in (3.2). This estimate assumes that the firm is maximizing its profits and that the firm can evaluate the residual demand it faces. As discussed in Section 2.1.1, even given these assumptions, these must be viewed as only approximate estimates of mark-up over competitive prices, since competitive prices at each bus m may deviate from the marginal costs $c'_m(q_m)$ at the market-clearing conditions. Moreover, to estimate the mark-ups above competitive prices more precisely, it is necessary to consider the generator capacity constraints, which will be discussed in the following two sections, 3.2.1.2 and 3.2.1.3.

3.2.1.2 Considering binding generator capacity constraints at market-clearing conditions

To consider the case of generators at their maximum capacity at market-clearing conditions, partition the generators owned by firm i into generators that are:

- at their maximum capacity, denoted by subscript f for fixed, and
- marginal, denoted by subscript v for variable.

The mark-up of each generator m owned by firm i that is marginal is still given by (3.1). Collect the mark-ups of the marginal generators owned by firm i into the vector Δp_v . Writing p_v and p_f for the vector functions representing the prices at the marginal and maximum capacity generators and using the symmetry of $\frac{\partial p}{\partial q}$, we have:

$$\Delta p_v = -\frac{\partial p_v}{\partial q}(q)q. \quad (3.3)$$

In contrast, the mark-up, Δp_f , of the generators at their maximum capacity will not satisfy the condition (3.1). Moreover, there would be non-zero “mark-up” of prices for such generators even with competitive offers since generators at their maximum capacity receive infra-marginal rents. This “mark-up” with competitive offers does not represent excess transfer above competitive levels. Nevertheless, above-competitive offers by the marginal generators owned by firm i would result in even higher mark-ups at the other generators

and we will estimate this effect to estimate the mark-up above competitive prices for the generators at maximum capacity. We will not attempt to estimate the higher mark-up of firm i 's generators at maximum capacity that is due to other firms' actions.

In particular, for the generators owned by firm i that are at maximum capacity, we will estimate the mark-up Δp_f of prices at their buses that is due to the mark-ups at the marginal generators owned by firm i . We will estimate Δp_f by considering the change Δq_v at the marginal generators of firm i that would result in changing the prices at those buses by Δp_v . By definition of the derivative, we have that:

$$\begin{pmatrix} \Delta p_f \\ \Delta p_v \end{pmatrix} \approx \begin{pmatrix} \frac{\partial p_f}{\partial q_v}(q) \\ \frac{\partial p_v}{\partial q_v}(q) \end{pmatrix} \Delta q_v. \quad (3.4)$$

To eliminate Δq_v from this expression, first assume that $\frac{\partial p_v}{\partial q_v}(q)$ is invertible and ignore capacity constraints of generators that are marginal at the market-clearing conditions. We have that:

$$\Delta p_f \approx \frac{\partial p_f}{\partial q_v}(q) \left[\frac{\partial p_v}{\partial q_v}(q) \right]^{-1} \Delta p_v.$$

Combining with (3.3), we obtain:

$$\begin{pmatrix} \Delta p_f \\ \Delta p_v \end{pmatrix} \approx - \begin{bmatrix} \frac{\partial p_f}{\partial q_v}(q) \left[\frac{\partial p_v}{\partial q_v}(q) \right]^{-1} \frac{\partial p_v}{\partial q_f}(q) & \frac{\partial p_f}{\partial q_v}(q) \\ \frac{\partial p_v}{\partial q_f}(q) & \frac{\partial p_v}{\partial q_v}(q) \end{bmatrix} q.$$

The estimated quantity-weighted average mark-up above competitive prices is then:

$$\frac{-q^\top \begin{bmatrix} \frac{\partial p_f}{\partial q_v}(q) \left[\frac{\partial p_v}{\partial q_v}(q) \right]^{-1} \frac{\partial p_v}{\partial q_f}(q) & \frac{\partial p_f}{\partial q_v}(q) \\ \frac{\partial p_v}{\partial q_f}(q) & \frac{\partial p_v}{\partial q_v}(q) \end{bmatrix} q}{\mathbf{1}^\top q}. \quad (3.5)$$

However, $\frac{\partial p_v}{\partial q_v}(q)$ can be singular. This occurs, for example, if multiple marginal generators are located at the same bus and therefore have the same shift factors with respect to all binding constraints, or if the number of marginal generators owned by firm i is greater than the number of binding constraints. In this case, we use the analogous expression to (3.5) that utilizes the pseudo-inverse of $\frac{\partial p_v}{\partial q_v}(q)$ [12, §A.5.4]. Using the pseudo-inverse results in a proxy to Δq_v that has the least Euclidean norm, and thus acts to under-estimate the profit-maximizing generation weighted mark-up.

Again, the TCMPI in (3.5) may also act to over-estimate the mark-up above competitive prices to the extent that competitive prices deviate from the marginal costs at the market-clearing conditions. Note that all generators are assumed to reveal their true capacity. In the particular case where all of firm i 's generators are at maximum capacity, we define the index to be zero. As mentioned above, this ignores the effect of other firms on the mark-up of firm i , which acts to under-estimate the mark-up above competitive prices in this case. On the other hand, by ignoring capacity constraints of generators that are marginal at the market-clearing condition, the index over-estimates

the actual mark-up over competitive conditions. We will consider the issue in the next section.

3.2.1.3 Considering binding generator capacity constraints at competitive conditions

In (3.4), Δq_v , the estimated generation output deviation from the case that firm i behaves competitively, does not consider the generation capacity constraints at the competitive condition:

$$\underline{q}_v \leq q_v - \Delta q_v \leq \bar{q}_v, \quad (3.6)$$

where $q_v - \Delta q_v$ represents the estimated production quantities at the competitive condition, and $\bar{q}_v, \underline{q}_v$ are the vectors of maximum and minimum generation capacities. Let V be the set of generators owned by firm i that are marginal at the market-clearing condition. Also define V^f and V^v as subsets of V which represent binding and marginal generators at the competitive condition, respectively. Note that $V^f, V^v \subseteq V$ and $V^f = V \setminus V^v$. Assuming sets V^f, V^v are known, we partition generators in V into generators that are:

- in set V^f , denoted by subscript vf ,
- in set V^v , denoted by subscript vv .

Collect the mark-ups of the generators in sets V^f and V^v into vectors Δp_{vf} and Δp_{vv} , respectively. Also, collect the production quantities deviated from the competitive condition of generators in set V^f and V^v into vectors Δq_{vf} and

Δq_{vv} , respectively. By definition of the derivative, we have that:

$$\begin{pmatrix} \Delta p_{vf} \\ \Delta p_{vv} \end{pmatrix} \approx \begin{bmatrix} \frac{\partial p_{vf}}{\partial q_{vf}}(q) & \frac{\partial p_{vf}}{\partial q_{vv}}(q) \\ \frac{\partial p_{vv}}{\partial q_{vf}}(q) & \frac{\partial p_{vv}}{\partial q_{vv}}(q) \end{bmatrix} \begin{pmatrix} \Delta q_{vf} \\ \Delta q_{vv} \end{pmatrix} \quad (3.7)$$

By definition of set V^f , we have

$$\Delta q_{vf} = q_{vf} - \bar{q}_{vf}. \quad (3.8)$$

Combining with (3.7) and re-arranging, we obtain:

$$\Delta q_{vv} \approx \left[\frac{\partial p_{vv}}{\partial q_{vv}}(q) \right]^{-1} \left(\Delta p_{vv} - \frac{\partial p_{vv}}{\partial q_{vf}}(q)(q_{vf} - \bar{q}_{vf}) \right), \quad (3.9)$$

where Δp_{vv} is the estimated price-cost mark-up given by (3.1). Using the symmetry of $\frac{\partial p}{\partial q}$, we have

$$\Delta p_{vv} = -\frac{\partial p_{vv}}{\partial q}(q)q. \quad (3.10)$$

Again, the pseudo-inverse can be utilized in the case that $\frac{\partial p_{vv}}{\partial q_{vv}}(q)$ is singular, as discussed in 3.2.1.2. Combining (3.7) and (3.9), the estimated mark-ups above competitive prices for generators in set V^f are:

$$\Delta p_{vf} \approx \frac{\partial p_{vf}}{\partial q_{vf}}(q)\Delta q_{vf} + \frac{\partial p_{vf}}{\partial q_{vv}}(q) \left[\frac{\partial p_{vv}}{\partial q_{vv}}(q) \right]^{-1} \left(\Delta p_{vv} - \frac{\partial p_{vv}}{\partial q_{vf}}(q)(q_{vf} - \bar{q}_{vf}) \right). \quad (3.11)$$

Furthermore, according to (3.4), the estimated mark-ups above competitive prices for generators that are binding at the market-clearing condition are:

$$\Delta p_f \approx \begin{bmatrix} \frac{\partial p_f}{\partial q_{vf}}(q) & \frac{\partial p_f}{\partial q_{vv}}(q) \end{bmatrix} \begin{pmatrix} \Delta q_{vf} \\ \Delta q_{vv} \end{pmatrix}, \quad (3.12)$$

where Δq_{vf} and Δq_{vv} are given by (3.8) and (3.9), respectively. The estimated quantity-weighted average mark-up above competitive price is then:

$$\frac{\begin{bmatrix} q_{vf}^\top & q_{vv}^\top & q_f^\top \end{bmatrix} \begin{pmatrix} \Delta p_{vf} \\ \Delta p_{vv} \\ \Delta p_f \end{pmatrix}}{\mathbf{1}^\top q} \quad (3.13)$$

Algorithm 1 Find binding generators at competitive conditions

Input: $\frac{\partial p_v}{\partial q_v}(q), q_v, \bar{q}_v$
Output: a conjectured V^f
 $V^f \leftarrow \emptyset, V^v \leftarrow V$
repeat
 $T \leftarrow \emptyset$
 Calculate Δq_{vv} using equation (3.9)
 for each i in V^v **do**
 if $q_i - \Delta q_i > \bar{q}_i$ **then**
 $T \leftarrow T \cup \{i\}$
 end if
 end for
 $V^f \leftarrow V^f \cup T, V^v \leftarrow V^v \setminus T$
until $V^f = V$ or $T = \emptyset$

The analysis above is based on the assumption that either set V^f or V^v is known. However, this information is usually not available a priori. Therefore, a procedure is developed to conjecture the set V^f . The basic idea is to iteratively calculate Δq_{vv} based on (3.9) and move generators violating capacity constraints from V^v to V^f , until the capacity constraints (3.6) are satisfied for all generators in V . Algorithm 1 describes the main steps of this procedure.

In this algorithm, at least one generator is added to set V^f at each iteration, or otherwise the algorithm terminates. Therefore, the above algorithm is guaranteed to terminate in no more than $|V|$ iterations. Minimum production constraints can also be included in the same manner to deal with the case that the estimated generation outputs at competitive conditions are lower than their minimum production levels. Note that this procedure only provides a conjecture about the set V^f , which might differ from the actual set of binding generators at the competitive condition. Also, since the estimated price-cost mark-ups for generators in V^v are used to approximate the mark-ups above competitive prices, the index might still over-estimate the mark-ups above competitive levels, particularly under the case that the derivatives of marginal costs are greater than zero.

3.2.2 Evaluation of market power index

The key to evaluating the TCMPI is to efficiently evaluate the JIRD $\frac{\partial p}{\partial q}$. Reference [72] describes the calculation of residual demand derivative when there are binding transmission constraints, using sensitivity analysis of market-clearing conditions, and [73] extends the methodology to calculate the JIRD matrix faced by a firm. Given the following information, the JIRD can be evaluated efficiently:¹

¹As observed in Chapter 2, in many markets, offers are piecewise-constant. The residual demand derivative implicitly assumes that the offer is differentiable. From a practical perspective, the transmission constrained inverse residual demand derivative is estimating the change in prices for a small change in injection. Consequently, the average slope of offers over a suitable small change in production is utilized, as discussed in Section 2.3.4.

- Binding transmission constraints and their shift factors for injection at each generator,
- Offer information, and
- Market-clearing quantities.

Once the JIRD is evaluated, we can evaluate the TCMPI based on (3.13).

3.2.3 Implementation

We have implemented the calculation of the TCMPI using the same framework as in Section 2.3. We use PowerWorld as the OPF solver and visualization tool, and implement the index calculation in MATLAB. Figure 2.2 shows the flow of data in an calculation. Sections 3.3.3 and 3.3.4 demonstrate the indices and their use for medium- and large-scale systems.

3.3 Case Studies of Firm-Based Transmission-Constrained Market Power Index

3.3.1 Single firm, no transmission constraints

This example is aimed to show the importance of considering binding capacity constraints at competitive conditions as discussed in 3.2.1.3. A single firm owning two generators, a and b , faces a residual demand $D - p$ MW, where D is a parameter characterizing overall demand level and p is the market-clearing price. The marginal costs of generator a and b are $10 + 0.1q_a$ \$/MWh and $10 + 0.2q_b$ \$/MWh, respectively, where q_a and q_b are the pro-

duction quantities. For various values of D , the market-clearing prices are evaluated for two cases:

- The generation firm maximizes its profit by selecting optimal production quantities, represented by the blue solid line in Figure 3.1, and
- The generation firm behaves competitively by using marginal costs as offers, represented by the green dashed line in Figure 3.1.

Note that the firm's profit-maximizing strategies are obtained using a Cournot model [68]. Without transmission constraints, we can easily obtain the optimal production quantities and corresponding prices.

Also, two indices are evaluated for each D based on the profit-maximizing case:

- The quantity-weighted average mark-ups considering binding capacity constraints at competitive conditions based on (3.13) (represented by the red dashed line in Figure 3.2), and
- The quantity-weighted average mark-ups ignoring capacity constraints at competitive conditions based on (3.5) (represented by the blue dashed line in Figure 3.2).

Note that in this particular example, the estimated quantity-weighted mark-up equals the estimated mark-up of either generator since no transmission constraints are presented. When $D > 230$, both generators become binding

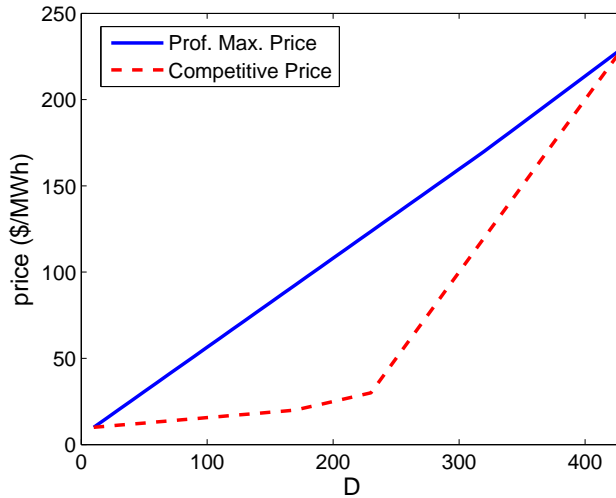


Figure 3.1: Market-clearing prices for the single firm, no transmission constraint example. The solid line (“Prof. Max. Price”) represents the prices when the generation firm maximizes its profit, while the dashed line (“Competitive Price”) represents the prices when the firm uses marginal costs as offers.

at the competitive condition. However, generator b is still marginal in the profit-maximizing case for $D > 230$. In this scenario, as shown in Figure 3.2, the estimated mark-up given by (3.5) significantly over-estimates the mark-up above the competitive price, while the estimated mark-up according to (3.13) corrects this inaccuracy and produces an estimate that is much closer to the actual mark-up above the competitive price obtained from a Cournot model.

3.3.2 Single firm in IEEE Reliability Test System

This example uses the IEEE 118-bus Reliability Test System to demonstrate that the TCMPI (3.13) can produce good estimates of the actual mark-ups above competitive prices. Assume there is a single firm maximizing its profit by choosing the optimal production quantities given all other sup-

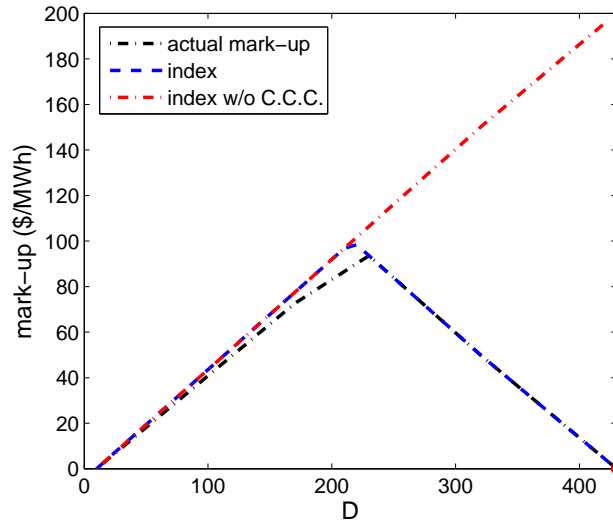


Figure 3.2: Actual and estimated mark-ups for single firm, no transmission constraint example. The blue line ('index') represents the estimated price mark-ups according to (3.5), and the red line ('index w/o C.C.C.') represents the estimated price mark-ups according to (3.13). This figure shows that considering capacity constraints at competitive conditions provide much more accurate estimates of actual mark-ups.

pliers' offers. The profit-maximizing strategies and corresponding market-clearing prices are computed using a mathematical program with equilibrium constraints (MPEC) [17]. Note that MPEC problems are typically non-convex and thus only guarantee solutions that are local optimal. Given the profit-maximizing production quantities of the firm and the corresponding market-clearing information, we compute the estimated mark-ups using (3.5) and (3.13). Table 3.1 shows the results, where q represents the profit-maximizing production quantities for the firm, q_{comp} represents the production quantities in the competitive condition, and \bar{q} represents the maximum capacities of each generator. The minimum capacities are all zero. The actual mark-ups are the profit-maximizing prices minus the competitive prices and are shown in the fifth column, with two estimates in the last two columns. As shown in the Table 3.1, the estimated mark-ups are reasonably good estimates of the actual mark-ups. The discrepancies in this example are presumably due to non-zero derivatives of marginal costs. That is, for marginal generators, the price-cost mark-ups over-estimate the mark-ups above competitive prices, as discussed in 3.2.1.3. Note that the estimated mark-ups given by (3.5), which ignore capacity constraints at the competitive condition, significantly over-estimate the actual mark-ups. Again, this example shows the need to incorporate capacity constraints at the competitive condition in order to produce more plausible estimates of mark-ups.

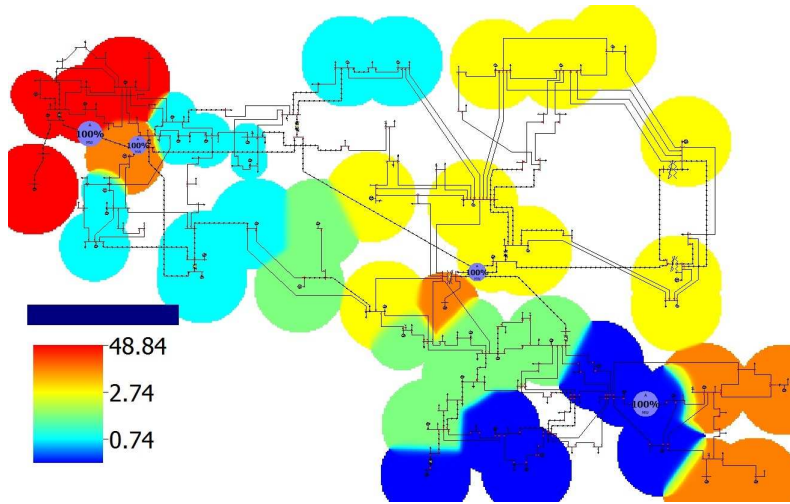


Figure 3.3: Contour map of firm-based transmission-constrained market power index for the IEEE 118-bus Reliability Test System. The units are \$/MWh. Generators owned by the same firm are assigned the TCMPI according to (3.13). This figure suggests that firm A has the ability to drive up the prices, presumably due one of the binding transmission constraints limiting power imports to the red contour areas. In contrast, the generator-based analysis fails to faithfully represent the market power owned by firm A.

Table 3.1: Actual and estimated mark-ups in IEEE 118-bus test system

Gen#	q	q_{comp}	\bar{q}	mark-up		
				actual	estimates ¹	estimates w/o c.c.c ²
1	77.75	105.23	162.5	6.69	9.98	9.98
2	58.52	78	78	0.32	-0.08	14.22
3	65	65	65	-0.21	-1.02	20.81
4	45.5	45.5	45.5	-0.38	-1.32	22.89
5	97.5	97.5	97.5	0.61	0.41	9.94
6	45.5	45.5	45.5	1.57	4.83	5.95
7	65	65	65	1.61	5.04	5.86
8	59.91	80.43	97.5	1.38	3.83	6.30
9	36.57	71.81	162.5	1.61	5.04	5.83
quantity-weighted average index				1.60	2.99	11.61

¹ Estimated mark-ups considering capacity constraints at competitive conditions.

² Estimated mark-ups ignoring capacity constraints at competitive conditions.

3.3.3 Multiple firms in IEEE Reliability Test System

Power flow data together with generic generator offers and artificial ownership data for the IEEE 118-bus Reliability Test System [54] are used for calculating the TCMPI (3.13). Note that the power flow data and generator offers are the same set of data as used in Section 2.4.1. This example demonstrates that the firm-based TCMPI is a more suitable tool for measuring market power in a transmission-constrained market, compared to the generator-based indices developed in Chapter 2. We assume the generators belong to six different owners, firms A,...,F, and each firm owns 5 to 13 generators.

Figure 3.3 shows the contours of the quantity-weighted average index

according to (3.13). All of the generators owned by the same firm are assigned the estimated quantity-weighted average for the firm and are displayed in the same color.² Note that there are four binding transmission constraints in the system, which are shown as the light blue pie charts.

The generators in the red contour areas are all owned by firm A, which has the largest estimated mark-up. This is presumably due to one of the binding transmission constraints limiting power imports to the red contour areas and making it less competitive. We will show in Section 3.5.2 that this firm does possess significant market power considering the effects of transmission constraints. All the other five firms have fairly small mark-ups and are all displayed in different colors.³

Comparing Figure 3.3 with Figure 2.6 in Section 2.4.1, it is clear that the generator-based market power indices proposed in Chapter 2 are not sufficient to analyze the behavior of a firm that owns generators at more than one location. In this case study, firm A, a pivotal firm as suggested in Section 3.5.2, has great potential to drive up the prices above the competitive levels. With the firm-based analysis proposed in this chapter, the estimated average mark-up of firm A is 48.84 \$/MWh, which indicates that firm A should be flagged as being of concern regarding exercising market power.

²The size of circles in the contours are determined by the parameter “Influence Region” in PowerWorld. There is no specific meaning on the size of circles of color contours.

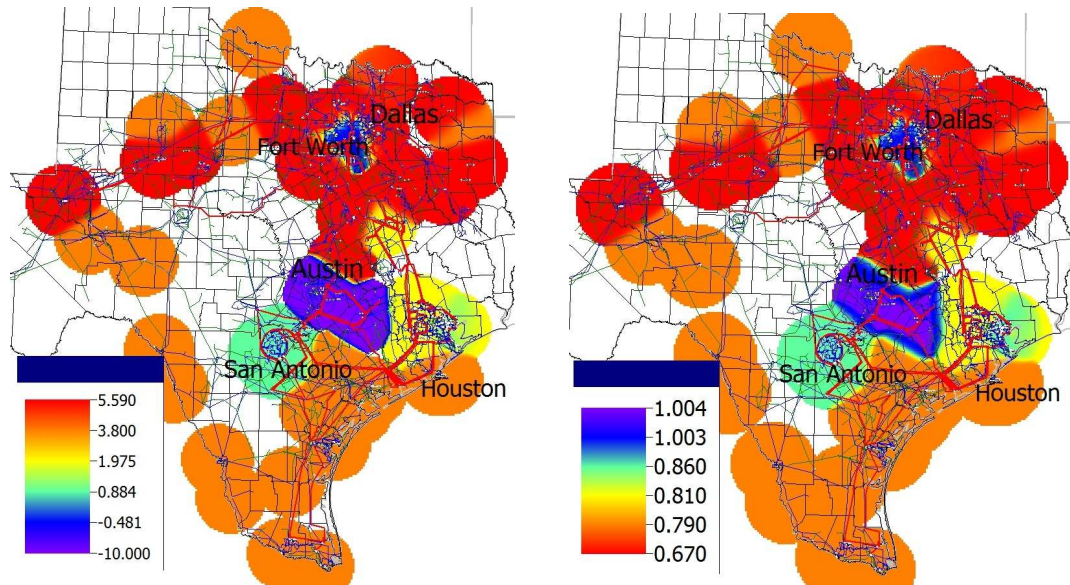
³There are several narrow stripes of yellow contours, such as the one between the aqua and orange contours in the middle of the figure. They are due to blending of colors and do not represent another firm.

In contrast, the individual generator analysis in Section 2.4.1 suggests that all generators owned by firm A have estimated mark-ups less than 3 \$/MWh, which are fairly small. In Section 2.4.1, the residual demand is based on all generation offers other than the generator to be analyzed. However, to analyze the profit-maximizing incentives for a firm, the offers from all the generators owned by the firm should be excluded from the residual demand calculation, since the firm can control the behavior of all its generators to maximize its profit. Therefore, the firm-based TCMPI is a more suitable tool to analyze market power of a firm.

3.3.4 ERCOT system

Power flow data together with generic generator offers⁴ and ownership and control data for the ERCOT system for summer peak demand are used for calculating the index. Note that the power flow data and generator offers are the same set of data as used in Section 2.4.2. We calculated the index based on security-constrained OPF (SCOPF) solution, assuming all generators are in-service. This example demonstrates that the firm-based TCMPI is a more suitable tool for measuring market power in a transmission-constrained market, compared to the generator-based indices developed in Chapter 2. There are 19 binding transmission constraints in the system, with three of them be-

⁴For instance, for a 600MW coal-fired generation unit, the offer curve, represented as a linear function, ranges from \$24-30/MWh. As another example, for a 400MW gas-fired steam turbine, the marginal offer, also represented as a linear function, ranges from \$63-71/MWh.



(a) Common log of transmission-constrained market power index in $\$/\text{MWh}$, except for purple contour, which has index of zero. (b) Transmission-constrained residual supply index (see Section 3.5.3).

Figure 3.4: Contour maps of two firm-based market power analyses for the ERCOT system. Six largest firms in terms of capacity share are displayed. This figure shows that the TCMPI and TCRSI are strongly correlated. However, this observation is not always true for other smaller firms in the system, which is discussed in Section 3.6.

ing pre-contingency constraints. Figure 3.4(a) shows the contours of estimated quantity-weighted average index according to (3.13). The contours represent the common log of the index. Again, all generators owned by a given firm are assigned the estimated quantity-weighted average for the firm and are displayed in the same color. For clarity, only the six largest firms in terms of capacity share are shown in the figure. The firm with the largest index is also the largest firm, which owns about 20% generation capacity. The generators owned or controlled by this firm are shown in red contours in the figure and are distributed in several areas in ERCOT. Note that the firm represented by purple contours has estimated mark-up of 0, since all of its generators are at full capacity. In the figure, we use a small number, 10^{-10} , to represent this mark-up, since common log of 0 is not defined.

The result shown in Figure 3.4(a) suggests that the market is much less competitive compared to the case in Section 2.4.2, where all generators are assumed to be owned by different firms. In Section 2.4.2, only five generators have estimated mark-ups higher than 1000 \$/MWh using generator-based market power indices. However, with the firm-based analysis, four firms, with 103 generators, have average estimated mark-ups higher than 1000 \$/MWh. We will show in Section 3.5.3 that these four firms indeed possess significant market power under transmission constraints. This example again demonstrates that the generator-based indices developed in Chapter 2 might underestimate the firm-based market power in a transmission-constrained market.

3.4 Transmission-Constrained Residual Supply Index

In this section, we consider another approach to firm-based analysis of market power. In the absence of binding transmission constraints, the residual supply index (RSI) of firm i is defined as follows:

$$\text{RSI} = \frac{\text{total available supply} - \text{available supply from firm } i}{\text{demand}}$$

The RSI measures the maximum available supply divided by the demand, without any supply from firm i . When the RSI is less than 1, firm i is said to be “pivotal” and has significant potential market power. On the other hand, a larger RSI value implies that firm i has less potential market power. In [59], empirical evidence of negative correlation between the RSI and price-cost mark-up was provided. This correlation indicates that the RSI might be a useful index to predict market power. Newbery [44] concludes that the RSI is particularly useful for the case of a single dominant firm or symmetric oligopoly.

However, the basic RSI definition ignores the impact of transmission constraints. In the presence of transmission constraints, a firm might be pivotal in particular geographic areas even if from a region-wide perspective (ignoring transmission constraints) there is no pivotal firm. CAISO uses an RSI-like approach [15] to assess the competitiveness of transmission constraints: it evaluates the extent to which a firm’s supply is necessary to produce counter-flow that relieves congestion on a particular constraint. In addition, CAISO applies “competitive path assessment” [14] to identify pivotal suppliers under trans-

mission constraints by repeatedly solving multiple OPFs that successively omit each given firm’s supply. This approach evaluates the shortfall in transmission capacity to meet demand and might be computationally expensive because of the need to solve an OPF for each firm.

In order to generalize the original definition of the RSI into the context of transmission constraints, but also provide a computationally efficient procedure, we developed a new index, the transmission-constrained residual supply index (TCRSI). Unlike the approaches adopted by CAISO, the TCRSI directly generalizes the original definition of the RSI while considering all constraints simultaneously. The TCRSI can measure the extent to which a firm’s supply is necessary to meet demand including the effects of transmission constraints. Generalizing the RSI, the TCRSI of firm i evaluates the residual supply at each load bus, in the absence of supply from firm i . As will be discussed in the next section, the TCRSI is represented as a single parameter that scales each load throughout the system conformally, and can be evaluated by solving a linear programming (LP) model if the transmission constraints are represented using DC power flow. The TCRSI can be computed efficiently by utilizing the dual simplex method, which will be further discussed in Section 3.4.3.

3.4.1 Definition of transmission-constrained residual supply index

Suppose that there are n generators, ℓ loads, and m transmission constraints. Variables $q \in \mathbb{R}^n$ represent the output quantity of each generator while $d \in \mathbb{R}^\ell$ is the vector composed of load quantities, which is given as

problem data. The matrix $H_g \in \mathbb{R}^{m \times n}$ consists of the shift factors for injection at each generator corresponding to each transmission constraint; similarly, $H_d \in \mathbb{R}^{m \times \ell}$ is the matrix of shift factors for injection at each load bus corresponding to each transmission constraint; and $b \in \mathbb{R}^m$ represents the limits for the transmission constraints. We assume that $\bar{q}_j, j = 1, \dots, n$, is the maximum available supply of each generator j . We again assume that K_i is the set of generators owned by firm i .

We define the TCRSI LP, which is used to calculate TCRSI for firm i , as follows:

$$\max_{q, t} t \tag{3.14a}$$

$$\text{subject to } \mathbf{1}^\top q - (\mathbf{1}^\top d)t = 0 \tag{3.14b}$$

$$H_g q - H_d(dt) \leq b \tag{3.14c}$$

$$0 \leq q_j \leq \bar{q}_j, j \notin K_i \tag{3.14d}$$

$$q_j = 0, j \in K_i \tag{3.14e}$$

where:

- (3.14b) is the power balance constraint;
- (3.14c) are transmission constraints;
- (3.14d) are maximum supply constraints;
- (3.14e) remove the supply of firm i .

With transmission constraints and maximum supply constraints, the above LP finds the maximum load $(\mathbf{1}^\top d)t$ that the system could meet after removing all available supply from firm i ; in other words, the LP maximizes the residual supply in the absence of firm i assuming demand is scaled conformally.

Definition 1. The TCRSI for the firm i is the optimal value of the linear program TCRSI LP, as specified in (3.14).

If the TCRSI is less than 1, it implies that there is insufficient supply to satisfy all the loads in the system without the supply from firm i ; that is to say, firm i is pivotal. When the TCRSI is greater than 1, the system has more than enough generation resources to satisfy all demand even without any supply from firm i . This observation is analogous to the interpretation of RSI.

Note that without the transmission constraints (3.14c), the optimal value of TCRSI LP evaluates the RSI. Therefore, the RSI can be interpreted as the optimal value of the relaxed TCRSI LP. Thus for a given firm i , the value of TCRSI is always less than or equal to the RSI. If a market monitor uses the TCRSI and RSI as screening tools for market power to flag the firms whose index is less than a given threshold, the firms that are flagged using the TCRSI would be a superset of firms flagged by the RSI. In other words, the TCRSI is a more comprehensive market screening tool since it flags all the firms that the RSI detects as pivotal, plus some firms that are pivotal due to transmission constraints that the RSI fails to discover.

The reason for using a single parameter to scale each individual load

conformally is that loads are distributed throughout the system with certain geographical patterns. For example, it is usually the case that metropolitan areas have significantly higher demand than rural areas. This pattern should be kept while we search for the maximum demand that the system could provide. In contrast, if each load is modeled as a free variable, the optimal solution might not represent the geographical distribution of the load in the system. This issue will be further discussed in Section 3.5.1 with reference to a small four bus example system.

3.4.2 Implications of transmission-constrained residual supply index

Unlike the TCMPI analysis, the TCRSI does not directly model the incentives for a firm to drive the prices above competitive levels. Nevertheless, as discussed in the first paragraph in Section 3.4, this type of *ex ante* analysis, which measures the extent to which a firm's supply is necessary to meet demand, might provide some useful insights concerning market power.

If the TCRSI of firm i is less than 1, it implies that the system cannot meet all demand without at least some supply from firm i . Assuming offers from other market participants are fixed, firm i could submit its generation offers at any arbitrarily high prices and at least some of its offered supply is guaranteed to be accepted by the ISO, given that the system has to satisfy all demand, all transmission constraints must be satisfied, and there exists no demand elasticity. As a result, a TCRSI value that is less than 1 implies an

absolute potential of exercising market power and should be flagged for further analysis.

In the case that the TCRSI of firm i is equal to or slightly greater than 1, the system can just meet all demand without the supply from firm i . However, the operator might need to dispatch some expensive generation resources. In this situation, firm i can offer strategically, resulting in its generators being dispatched at prices much higher than competitive levels.

3.4.3 Implementation

Section 3.4.1 defined the TCRSI of a given firm, which is derived from solving a LP problem. To calculate the TCRSIs for all the market participants, multiple TCRSI LPs have to be solved. The TCRSI LP problem (3.14) has a similar structure to an optimal power flow problem (OPF), except for the difference in the objective function. Solving an OPF is computationally expensive if the system size is large. Therefore, an efficient implementation is desired.

The only changes between the TCRSI LP formulation (3.14) for different firms are changes in the generators' maximum available supply. More specifically, a different set of generators' maximum available supplies are suppressed to 0 when formulating the TCRSI LP for each different market participant. This characteristic can be exploited by using the dual simplex method to solve all the TCRSI LPs starting from a dual feasible solution. Once the TCRSI LP is solved for a given firm, we keep the basis of the current LP, change

the right-hand sides of the constraints and continue the dual simplex method until optimality conditions are reached. This technique is called “warm start” and is supported by most LP solvers⁵.

We used Gurobi 2.0 as the LP solver and wrote the code in the Python language to calculate the TCRSI. A small four bus example is first developed to provide a comparison between RSI and TCRSI. Then the TCRSI is also evaluated for the two test cases from Sections 3.3.3 and 3.3.4. Table 3.2 shows the results of computational performance for these two test cases. We observe that the computational efficiency for solving a large system is significantly improved by using warm start, compared to solving TCRSI LP from scratch for each firm, labeled “Cold start” in Table 3.2.

3.5 Case Studies of the TCRSI

3.5.1 4 Bus System

We use a simple four bus system to illustrate why the TCRSI is a more suitable index than the RSI in the context of transmission constraints. Figure 3.5 shows the one-line diagram of the four bus system. There are three generators and two loads in the network. Suppose that each generator is owned by a different firm and the subscripts of generators and loads also represent the bus where they are located. Assume that the transmission line from bus 2 to 3 has a thermal limit of 300MW, while the other lines have sufficiently

⁵In principle, the same approach could also be used to speed up calculations in the CAISO “competitive path assessment” approach.

Table 3.2: Computational performance

Test case	Method	CPU Time ¹	Iterations ²
118-bus RTS ³	Cold start	0.07	419
	Warm start	0.06	417
ERCOT ⁴	Cold start	180.3	38983
	Warm start	41.5	11310

¹ The units are second. Based on 3.0 GHz CPU, 16GB RAM Linux workstation. Note that for the ERCOT case, the CPU time does not include the time for solving SCOPF.

² Number of total simplex iterations.

³ The Reliability Test System has 118 buses, 54 generators, 91 loads, and 194 transmission constraints.

⁴ The ERCOT system has 5526 buses, 3518 loads, 564 generators, and 6836 transmission constraints.

Table 3.3: TCRSI and RSI of the four bus system

Firm	TCRSI	RSI
g_1	1	1
g_2	1.14	1.14
g_3	0.58	1

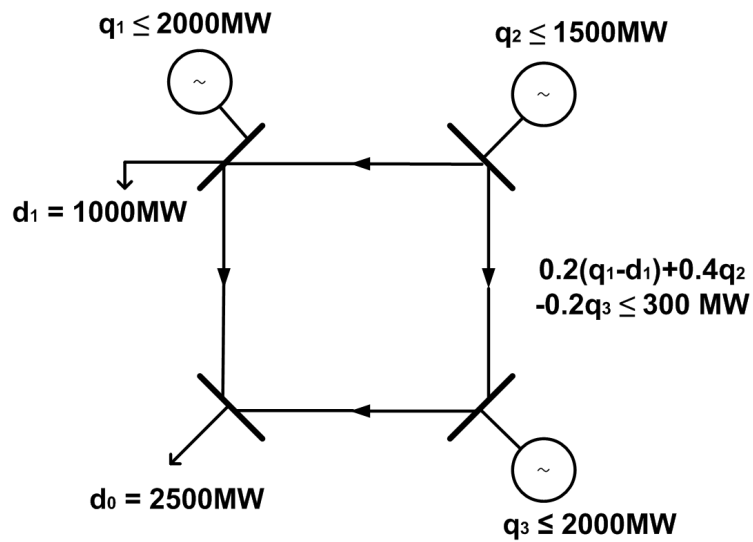


Figure 3.5: One-line diagram of the four bus system. The RSI analysis implies that there is no pivotal firm. However, the TCRSI of generator 3 is 0.58 and indicates that generator 3 is a pivotal supplier. The TCRSI reveals that generator 3 has great potential for exercising market power since it can offer its production at arbitrarily high prices and consequently drive up the market-clearing prices.

large thermal limits so that the associated transmission constraints are never binding.

Table 3.3 shows the TCRSI and RSI of each generator. The RSI analysis implies that there is no pivotal firm, since the RSI of each firm is greater than or equal to 1. However, the TCRSI of generator g_3 is 0.58 and indicates that g_3 is a pivotal supplier. Without the supply from g_3 , the maximum conformal loads the system could serve are 1458.3MW at d_0 and 583.3MW at d_1 , due to the congested line from bus 2 to bus 3.

To further motivate the TCRSI index as defined in (3.14), consider the following alternative formulation. Instead of scaling load conformally as in (3.14b), suppose that we had modeled each load as an independent variable and then maximized the total demand without the supply of each given firm. In the case of omitting the supply from generator g_3 , the resulting loads at buses 0 and 1 would be $d_0 = 0\text{MW}$ and $d_1 = 3500\text{MW}$, respectively. This alternative modeling distorts the original geographical pattern of demand and would give the false impression that g_3 were not pivotal. In this example, generator g_3 is in fact pivotal due to its particular situation with respect to transmission constraints. The TCRSI reveals that g_3 has great potential for exercising market power since it can offer its production at arbitrarily high prices and consequently drive up the market-clearing prices.

Table 3.4: TCRSI and RSI of the IEEE 118-bus Reliability Test system

Firm	TCRSI	RSI	Capacity Share ¹
A	0.61	3.30	0.09
B	1.00	3.10	0.14
C	0.85	2.83	0.22
D	1.48	3.17	0.13
E	1.23	3.06	0.15
F	1.39	2.76	0.24

¹ defined as $\frac{\text{firm's total capacity}}{\text{total capacity in the system}}$.

3.5.2 IEEE 118-bus Reliability Test System

We evaluated the TCRSI for the IEEE 118-bus Reliability Test System with the same data set used in Section 3.3.3. Table 3.4 shows the TCRSI, the RSI, and the capacity share of each firm. We assume the available supply of each generator is equal to its capacity. Ignoring transmission constraints, the system has an abundance of capacity to satisfy demand since the RSI of all firms are higher than 2.5. According to the RSI, there are no pivotal firms in the system. However, the TCRSI indicates that firms A and C are actually pivotal considering transmission constraints. Interestingly, firm A, which has the smallest TCRSI value, also has the largest RSI value and the least capacity share. This example again demonstrates that RSI analysis is insufficient to indicate the pivotal firms in the context of transmission constraints. Note that firm A also has largest estimated mark-up according to the analysis in Section 3.3.3.

3.5.3 ERCOT system

We evaluated the TCRSI for the ERCOT system with the same data set used in Section 3.3.4 and assuming that the capacity of each generator represents its available supply. In addition to all the pre-contingency transmission constraints, the binding and near binding post-contingency constraints from the SCOPF results are also incorporated into the analysis.⁶ Table 3.5 lists the six firms with the largest capacity share. The RSI analysis indicates that only one firm is pivotal, while the TCRSI reveals that four market participants are pivotal firms due to the effect of transmission constraints. Note that firms B and D have a relatively small capacity share, compared to firms A and C. However, the location of generators owned by firms B and D causes them to be pivotal suppliers.

Figure 3.4(b) in page 76 shows the contour map of the TCRSI for the ERCOT system. In both Figures 3.4(a) and 3.4(b), the generators owned by firm A are shown in red contours: firm A has the smallest TCRSI and also the largest estimated mark-up among all the firms. We can also observe that among the six selected firms, the firms with larger index according to (3.5) also have lower TCRSI. However, this observation is not always true for other firms in the system, which will be further discussed in Section 3.6.

⁶That is, we potentially omit some actually binding post-contingency constraints for some firms. This means that the values we obtain may over-estimate the actual TCRSI.

Table 3.5: TCRSI and RSI for selected firms in the ERCOT System

Firm	TCRSI	RSI	Capacity Share
A	0.668	0.976	0.199
B	0.793	1.151	0.056
C	0.809	1.077	0.116
D	0.859	1.158	0.050
E	1.003	1.184	0.028
F	1.004	1.184	0.029

3.6 Comparison of Indices

Two different categories of firm-based transmission-constrained market power assessment tools have been proposed in this chapter. The TCMPI involves *ex post* analysis, that is, it depends on offer information and the OPF results. Based on this information, the index estimates the quantity-weighted average mark-up above competitive levels for a given firm. A firm should be flagged as of potential concern if its index value is high.

In contrast, the TCRSI is an *ex ante* analysis, which does not require generation offers nor the OPF results, and estimates the extent to which a firm’s supply is necessary to meet the demand under transmission constraints. Note that the TCRSI does not depend on the behavior of the market participants and is aimed at predicting the potential for market power beforehand.

As a further distinction between these approaches, note that the transmission-constrained market power index is a “small signal” analysis; that is, it is based on the derivative of market-clearing price with respect to demand. On the

other hand, the TCRSI is a “large signal” analysis, since all the available supply from a firm is withdrawn during the analysis.

Although these two market power analysis schemes are different from several perspectives, they have one common characteristic: they both focus on what a firm could do, given the behavior and characteristics of the others. The two different approaches have different applications in evaluating market power. The TCMPI is suitable for detecting the exercise of market power, since it is based on the hypothesis that a firm is maximizing its profit. In other words, if a firm is maximizing its profit and significantly raising the prices above competitive levels, the index would reflect this behavior.

On the other hand, the TCRSI is an appropriate tool for predicting the possession of market power. It does not require generation offers and market clearing results but could identify the firms that have absolute market power. Therefore, both tools can be integrated into the market power analysis flow due to their different implications for market power. Figure 3.6 shows a scatter plot of the TCRSI versus the TCMPI (on a common log scale) for the ERCOT system. From the figure, we can observe that the TCRSI and TCMPI are complementary in some cases: some firms are pivotal but their estimated mark-ups are just moderately large, while some firms have very large estimated mark-ups even without being pivotal. The numerical results are consistent with our claim that these two approaches are complementary in assessment of market power and can both be utilized to analyze the behavior of the firms.

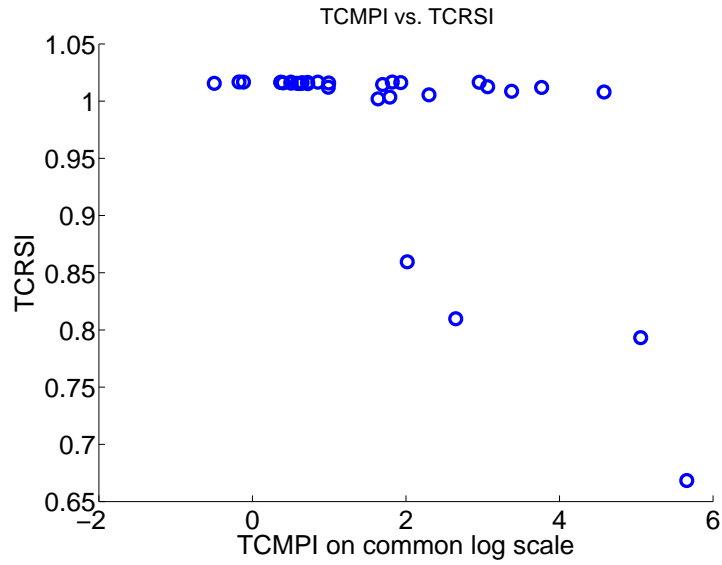


Figure 3.6: TCRSI versus TCMPI for the ERCOT system. This figure shows that TCRSI and TCMPI are complementary in some cases: some firms are pivotal but their estimated mark-ups are just moderately large, while some firms have very large estimated mark-ups even without being pivotal. This results suggest that these two approaches are complementary in assessment of market power and can both be utilized to analyze the behavior of the firms.

3.7 Conclusion

Market power continues to be a problematic issue that causes electricity markets inefficient. Unfortunately, most existing approaches of analyzing market power in transmission-constrained markets only have tenuous connections to economic incentives of exercising market power. In this chapter, we propose two different approaches to firm-based market power analysis: the transmission-constrained market power index (TCMPI), and the transmission-constrained residual supply index (TCRSI). We not only show the rationale of these approaches but also develop computationally efficient implementations for them, which can cope with large-scale systems such as the ERCOT system. Numerical results show that the TCMPI can indeed provide reliable estimates on price mark-ups above competitive levels and thus help the market monitor restore the prices to appropriate levels. On the other hand, we also demonstrate that the TCRSI provides more useful insights on market power in transmission-constrained markets, compared to the RSI. The two proposed indices are compared with qualitative analysis together with experimental results; we conclude that both methods should be incorporated into market power analysis flow, because their different insights about market power can help market monitor better understand the strategic behavior of market participants.

Chapter 4

Frequency-Constrained Stochastic Economic Dispatch

4.1 Introduction

Electric power grid operations need reserves to balance supply and demand under system contingencies, and appropriate reserve allocation is thus essential for efficient system and market operations. In Section 1.3, we summarize the three issues regarding reserves that should be addressed carefully, they are,

- variability of wind,
- geographical allocation of reserves, and
- sufficient primary frequency control.

Failing to incorporate these issues in reserve allocation may lead to unreliable, inefficient system operations. Even without significant levels of wind power integration, system operators have growing concerns about geographical allocation of reserves [29] [74] and primary frequency control [36] [45]. This chapter proposes a new model, the stochastic economic dispatch model (SED), to find the optimal schedules for energy and reserves that result in the

minimum expected operation costs. The three issues regarding reserves are systematically addressed in the proposed model.

Stochastic programming (or stochastic optimization) is a promising approach to dispatch generation and allocate reserves under uncertainty. Takriti and Birge [65] describe a multi-stage stochastic program to solve a unit commitment problem with uncertain demand, but without transmission constraints. Bouffard *et al.* [11] propose a stochastic market-clearing model considering transmission constraints. Their model deals with equipment failures and uses the expected value of unserved energy as the security metric. Ruiz *et al.* [56] propose a two-stage stochastic unit commitment model that incorporates reserve requirements. Kimball *et al.* [35] propose a stochastic optimal power flow model and solve it by Benders' decomposition. They demonstrate that a stochastic model can indeed reduce the expected operation costs. Saric *et al.* [58] formulate the single period market clearing process using a two-stage stochastic program and show that the model is computationally tractable using an interior-point algorithm together with Benders' decomposition. Note that none of the models reported in these references represent the response of system frequency. In other words, these model use the "steady-state" equilibrium to represent the post-contingency operation. However, according to [36], the quality of frequency control in U.S. has been declining in recent years and thus primary frequency control should be incorporated in the operation planning process. This observation motivates us to propose a new model that characterize post-contingency frequency dynamics in order to ensure sufficient

primary frequency control.

Several papers in the literature attempt to incorporate frequency response into the dispatch problem. Somuah and Schweppe [62] develop an economic dispatch model with a constraint on the maximum frequency deviation given a specified disturbance. The frequency constraint incorporates turbine dynamics, which makes the problem non-convex.¹ Restrepo and Galiana [55] develop a unit commitment model with requirements for primary frequency control. They characterize the amount of primary frequency control a unit could provide using the steady-state frequency deviation, without consideration of under-frequency load shedding (UFLS). Doherty *et al.* [18] develop a market-clearing model with linear frequency control constraints. They simulated a large number of events using a simplified dynamic model to construct a feasible region in terms of system kinetic energy, reserve procurement, and size of contingency. Ruiz and Sauer [57] describe a simplified dynamic model to estimate the amount of UFLS and construct a demand curve for reserves. The models in both [18] and [57] aggregate the primary frequency control of each unit into a single quantity. That is, each unit's impact on the system frequency, depending on its response speed and operating point, is not well differentiated.

¹Although the authors of [62] claim the frequency deviation function is convex, no rigorous proof is provided. In fact, the co-existence of the fifth and sixth term in equation (13) in [62] makes the function non-convex.

This chapter proposes a new model, the stochastic economic dispatch model (SED), to find the optimal schedules for energy and reserves that minimize expected operation costs (including penalties for load shedding). Our main contributions are summarized as follows:

- We propose the idea of incorporating minimum frequency constraints in a stochastic optimization framework, so that primary frequency control and geographical allocation of reserves can be co-optimized to minimize the expected operation costs.
- We propose a new formulation for frequency constraints and a method for estimating the unserved demand resulting from UFLS. To improve the accuracy of the frequency constraints, we characterize each individual generator’s impact on the system frequency based on its maximum ramp rate and “headroom,” i.e., available capacity left for increasing output. Moreover, the proposed formulation is convex and computationally efficient. Comparison with transient simulation is also provided to show the accuracy of our formulation.
- We apply the L-shaped decomposition method [61] for solution and provide numerical results based on medium- and large-scale systems. Computational efficiency of the formulation is demonstrated. We also show that the proposed model improves the dispatch in terms of expected operation costs, compared to deterministic methods.

The rest of this chapter is organized as follows: Section 4.2 introduces the formulation of SED. Section 4.4 presents the decomposition algorithm. Section 4.5 provides numerical examples, and Section 4.6 concludes.

4.2 Stochastic Economic Dispatch Model

4.2.1 Introduction

The proposed model, the stochastic economic dispatch model (SED), is formulated as a two-stage stochastic convex program. The first-stage decisions represent pre-contingency generation dispatch and procurements of reserves. After knowing a specific realization of system conditions, second-stage decisions, the generation re-dispatch decisions, are made to optimally respond to this specific scenario. Note that the feasible region of second-stage decisions are influenced by the first-stage decisions. For example, if more reserves are procured in the first-stage, the more generation resources would be available in the second-stage. The objective of SED is to minimize the expected operation costs, subject to both pre- and post-contingency operation constraints over an operational time frame.

In SED, a contingency scenario consists of one or more generator failures, together with significant demand forecast or wind forecast errors. Note that in the second-stage, all unknown information in the first-stage is revealed. That is, the availability of each generator and the exact value of demand and wind power output are known. We assume that any generation outages occur at the beginning of the operational time frame. The operational time frame is

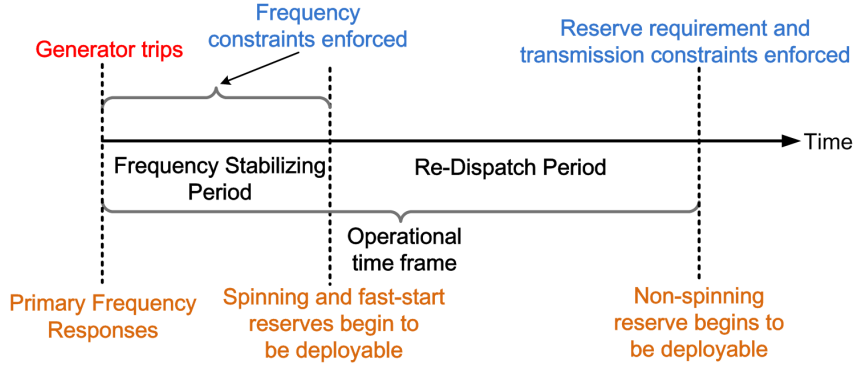


Figure 4.1: Timeline of post-contingency operation represented by the second-stage problem. In the frequency stabilizing period, system frequency is stabilized by primary frequency responses of generators together with UFLS. In the re-dispatch period, the system operator re-dispatches generation to restore the system security.

divided into two non-overlapping periods, as shown in Figure 4.1. These periods characterize the typical process of restoring the system back to a normal state after the disturbance:

- **Frequency stabilizing period (FSP):** Over a duration of t^S seconds, system frequency is stabilized by primary frequency responses of generators together with UFLS, if necessary. Transmission constraints are ignored in this period since the main operational goal here is to stabilize frequency; also, a short period of violation of transmission constraints usually does not threaten system reliability. The FSP is analogous to the “emergency state” in [19], where the control objective is to relieve system distress and prevent further degradation.
- **Re-dispatch period (RDP):** After the FSP, the system operator re-dispatches generation to restore the system security. Involuntary load

shedding is applied if generators alone are insufficient to restore system security. We suppose that spinning reserve and fast-start reserves are ready to be deployed in this period.² At the end of the RDP, non-spinning reserves can be brought on-line to meet the minimum spinning reserve requirement. The typical duration of the RDP is from 15 to 30 minutes [48]. This period is analogous to the “restorative state” in [19], where the control objective is the safe transition from partial to full satisfaction of all security constraints.

Unlike other stochastic dispatch models in the literature [35] [58], which use the post-contingency “steady-state” equilibrium to characterize the post-contingency operation, SED represents the process of disturbance recovery by incorporating frequency dynamics after the disturbance. Figure 4.1 shows the timeline for the post-contingency operation described above. It illustrates the timing of enforcing different operational constraints and the availability of each type of reserve.

4.2.2 Notation

This section summarizes the notation used in the development of the model. To simplify the presentation of the model but without loss of generality, we assume that each bus has exactly one generator and one load. Also, wind power is modeled as negative load in our formulation and “net load” is

²Spinning reserves are usually available prior to the beginning of the RDP.

defined as load minus wind power output.³

Indices and Sets:

$i \in \mathcal{N}$ buses

$i \in \mathcal{N}^W$ subset of buses with wind power generation

$i \in \mathcal{N}^G$ subset of buses with generators equipped with governors

$\ell \in \mathcal{L}$ transmission constraints

$\omega \in \Omega$ contingency scenarios

Random parameters:

$\tilde{d}_{i\omega}$ net load at bus i at the end of the RDP under scenarios ω [MW]

$\tilde{a}_{i\omega}$ availability of generator at bus i under scenarios ω : 1 for normal operation,
0 for outage

Parameters:

π_ω probability of contingency scenario ω

π_0 probability that no contingencies occur

³The wind generation resources could also be modeled as non-dispatchable generators with equivalent inertia.

- t^S length of frequency stabilizing period [s]
- t^L duration of involuntary load shedding until restored [h]
- d_i net load forecast at bus i [MW]
- b_ℓ limit of transmission constraint ℓ [MW]
- $A_{\ell i}$ shift factors of injection at bus i on constraint ℓ
- u_i commitment status of generator at bus i : 1 is on-line, 0 is not
- α_i for generator at bus i : 1 is fast start capable, 0 is not
- H_i inertia constant⁴ for generator at bus i [s]
- \bar{g}_i maximum output of generator at bus i [MW]
- \underline{g}_i minimum output of generator at bus i [MW]
- \bar{g}_i^{SP} spinning reserve limit of generator at bus i [MW]
- ρ_i upper ramp rate limit of generator at bus i [MW/s]
- c penalty cost for unserved load [\$/MWh]
- s post-contingency spinning reserves requirement [MW]
- v_i fraction of UFLS deployed at bus i

⁴ $H = \frac{\text{stored kinetic energy at nominal frequency}}{\text{machine MVA rating}}$

First-stage decision variables:

x_i generation dispatch of generator at bus i [MW]

r_i^{SP} spinning reserve from generator at bus i [MW]

r_i^{FS} fast-start reserve from generator at bus i [MW]

r_i^{NS} non-spinning reserve from generator at bus i [MW]

Second-stage decision variables:

$x_{i\omega}^{\text{R}}$ generation re-dispatch at bus i during the RDP under scenario ω [MW]

l_{ω}^{S} load shedding triggered in the FSP under scenario ω [MW]

$l_{i\omega}^{\text{R}}$ load shedding triggered in the RDP at bus i under scenario ω [MW]

$y_{i\omega}$ restored spinning reserve from generator at bus i under scenario ω [MW]

Functions:

C_e energy cost incurred in operational time frame [\$]

C_r reserve cost incurred in operational time frame [\$]

Q_{ω} second-stage post-contingency cost in operational time frame [\$] (see Section 4.2.3 and Section 4.3.2)

4.2.3 Stochastic economic dispatch model

Let $x \in \mathbb{R}^{|\mathcal{N}|}$ be the vector collecting together the dispatch variables x_i , $i \in \mathcal{N}$. Similarly, $r \in \mathbb{R}^{3|\mathcal{N}|}$ is the collection of all reserves variables, r_i^{SP} , r_i^{FS} , and r_i^{NS} , $i \in \mathcal{N}$. Function Q_ω , the post-contingency cost, will be introduced in Sections 4.2.4 and 4.3. SED is formulated as:

$$\min_{x,r} \quad \pi_0 (C_e(x) + C_r(r)) + \sum_{\omega \in \Omega} \pi_\omega Q_\omega(x, r) \quad (4.1a)$$

$$\text{s.t.} \quad \sum_{i \in \mathcal{N}} (x_i - d_i) = 0 \quad (4.1b)$$

$$\sum_{i \in \mathcal{N}} A_{\ell i} (x_i - d_i) \leq b_\ell, \forall \ell \in \mathcal{L} \quad (4.1c)$$

$$x_i + r_i^{\text{SP}} \leq u_i \bar{g}_i, \forall i \in \mathcal{N} \quad (4.1d)$$

$$x_i \geq u_i \underline{g}_i, \forall i \in \mathcal{N} \quad (4.1e)$$

$$r_i^{\text{SP}} \leq u_i \bar{g}_i^{\text{SP}}, \forall i \in \mathcal{N} \quad (4.1f)$$

$$r_i^{\text{FS}} \leq (1 - u_i) (\alpha_i) \bar{g}_i, \forall i \in \mathcal{N} \quad (4.1g)$$

$$r_i^{\text{FS}} + r_i^{\text{NS}} \leq (1 - u_i) \bar{g}_i, \forall i \in \mathcal{N}, \quad (4.1h)$$

The objective of SED (4.1a) is to minimize the expected operation costs. The pre-contingency operation costs consist of energy production costs and reserve procurement costs, weighted by π_0 , the probability of no occurrence of contingency. The post-contingency costs Q_ω , the key innovation of SED, includes the penalty costs for unserved load and re-dispatch energy costs. We assume that function C_e is convex quadratic and that C_r is linear.⁵ Constraint (4.1b)

⁵Typical generation costs can be well approximated by convex quadratic functions [71, §2], while reserve costs are modeled as linear functions in most electricity markets [21, §4.4.7] [40, §4.2].

is the power balance constraint that ensures supply meets demand. Constraints (4.1c) are transmission constraints that limit the flow on each transmission line, where $A_{\ell i}$ represents the amount of power flowing on line ℓ given a unit injection at bus i and withdrawal at the reference bus. We use DC power flow here, where $A_{\ell i}$ can be derived from linearization about the real-time operating point [71]. Note that $N - 1$ transmission security constraints are imposed to ensure that there will be no line-flow violations if any single line is removed from the system. Constraints (4.1d) and (4.1e) place lower and upper limits, respectively, on power generation for each generator. Constraints (4.1f) are maximum procurable spinning reserve constraints, which are typically determined by the ramp rate limit. Constraints (4.1g) specify that only fast-start capable generators can provide fast-start reserves, while constraints (4.1h) indicate that off-line capacity can provide either fast-start or non-spinning reserve, but not both. Note that the primary frequency control at bus i is implicitly determined by the generation dispatch x_i , which decides the amount of available capacity left for increasing output in the contingency state. We do not model regulation reserves here, but they can also be incorporated.

4.2.4 The second-stage problem

The second-stage problem models post-contingency system operations. The amount of unserved demand due to UFLS in the FSP is estimated and penalized. Directly after the FSP, the system operator re-dispatches generation

in order to restore system security. The objective in the RDP is to minimize the generation costs and the penalty costs of involuntary load shedding. In practice, the system operator might run a sequence of transmission-constrained economic dispatches as load varies with time in the RDP. Because of computational limitation, only a single re-dispatch is modeled here, which is based on the assumed known net load realization at the end of the the RDP. Let $x_\omega^R \in \mathbb{R}^n$ be the vector collecting the dispatch variables x_i . Function $\hat{p}_\omega^S(x)$, which is developed in Section 4.3, represents the amount of unserved load resulting from UFLS. The second-stage problem is formulated as follows:

$$Q_\omega(x, r) \tag{4.2a}$$

$$\equiv \min_{x_\omega^R, l_\omega^S, l_\omega^R} C_e(x_\omega^R) + ct^L(l_\omega^S + \sum_{i \in \mathcal{N}} l_{i\omega}^R) \tag{4.2a}$$

$$\text{s.t. } l_\omega^S \geq \hat{p}_\omega^S(x) \tag{4.2b}$$

$$l_\omega^S + \sum_{i \in \mathcal{N}} (x_{i\omega}^R + l_{i\omega}^R - \tilde{d}_{i\omega}) = 0 \tag{4.2c}$$

$$\sum_{i \in \mathcal{N}} A_{\ell i} (x_{i\omega}^R + v_i l_\omega^S + l_{i\omega}^R - \tilde{d}_{i\omega}) \leq b_\ell, \forall \ell \in \mathcal{L} \tag{4.2d}$$

$$y_{i\omega} \leq \tilde{a}_{i\omega} (x_i + r_i^{\text{SP}} + r_i^{\text{FS}} + r_i^{\text{NS}} - x_{i\omega}^R), \forall i \in \mathcal{N} \tag{4.2e}$$

$$y_{i\omega} \leq \bar{g}_i^{\text{SP}}, \forall i \in \mathcal{N} \tag{4.2f}$$

$$\sum_{i \in \mathcal{N}} y_{i\omega} \geq s \tag{4.2g}$$

$$\tilde{a}_{i\omega} \underline{g}_i \leq x_{i\omega}^R \leq \tilde{a}_{i\omega} (x_i + r_i^{\text{SP}} + r_i^{\text{FS}}), \forall i \in \mathcal{N} \tag{4.2h}$$

$$\tilde{d}_{i\omega} \geq l_{i\omega}^R \geq 0, \forall i \in \{\mathcal{N} \setminus \mathcal{N}^W\}, \tag{4.2i}$$

Note that we ignore the start-up and min-load costs of deploying fast-start and non-spinning reserves. The parameter c represents the penalty cost of

involuntary load shedding. Constraint (4.2b) places a lower limit on the unserved load due to UFLS. After the FSP, the transmission constraints (4.2d) are enforced to prevent over-loaded lines and restore transmission security. Note that the portion of load that is disconnected in the FSP are considered in constraints (4.2c) and (4.2d). Constraints (4.2e)-(4.2g) ensure that spinning reserve is restored to the minimum level to meet the reliability standard [48, §R6], assuming that non-spinning reserves are on-line at the end of the RDP. Constraints (4.2h) enforce upper limits on power generation for each generator in the RDP, assuming that only spinning and fast-start reserves can provide energy during the RDP. Note that we assume that non-spinning reserves are only deployable at the end of the RDP for restoring reserves. That is, they cannot provide energy during the RDP.⁶ Constraints (4.2i) limit load shedding at each bus to the total load at that bus.

The key innovation of SED is to formulate $\hat{p}_\omega^s(x)$ as a convex function. According to proposition 2.1 and 2.3 in [22], Q_ω is a convex function in (x, r) if $\hat{p}_\omega^s(x)$ is a convex function. The convexity of Q_ω further ensures SED a convex optimization problem.

4.3 Estimating Unserved Load Due to UFLS

This section describes a method to estimate the amount of UFLS that must be applied, given minimum frequency requirements, and incorporates

⁶We therefore ignore the minimum production limits of non-spinning reserves.

that estimate into the second-stage problem (4.2). The estimated involuntary load shedding $\hat{p}_\omega^s(x)$ is formulated as a convex function of the generation dispatch vector x . Note that convexity is a crucial property for applying decomposition methods and ensuring global optimality. We also provide analytical expressions for evaluating the subgradients of $\hat{p}_\omega^s(x)$, in order to efficiently generate linear approximations for the decomposition algorithm.

Directly after a disturbance, the frequency of the system declines and generators with governors respond by increasing mechanical power output in order to stabilize system frequency. If primary frequency responses from generators alone cannot maintain the frequency above a threshold, under-frequency load shedding (UFLS) will be triggered to protect the system from further damage. We assume that following a disturbance, all generators with headroom ramp their mechanical power production at the maximum ramp rate limits. This is a standard assumption in simplified frequency dynamic models [57] [62].

4.3.1 Evaluating system frequency

We can estimate system frequency at an instant in time by evaluating the stored kinetic energy after the disturbance. Recall that the stored kinetic energy E of rotating mass is

$$E(f) = \frac{1}{2}I(2\pi f)^2,$$

where I is the system inertia and f is the system frequency in Hz. Since f is non-negative, we can express frequency as a function of E :

$$f(E) = \frac{1}{2\pi} \sqrt{\frac{2E}{I}}. \quad (4.3)$$

Because f is a monotonic function in E , enforcing the minimum frequency requirement would be equivalent to enforcing a lower bound on stored kinetic energy. Assuming that the system is at the nominal frequency before the disturbance and no UFLS is applied, we obtain the frequency constraints

$$E_\omega + \Delta e_\omega(x, t) \geq \underline{E}_\omega, t \in [0, t^s], \quad (4.4)$$

where:

- $\Delta e_\omega(x, t)$ is the change in the stored kinetic energy at time t , given pre-contingency dispatch x .
- $E_\omega = \sum_{i \in \mathcal{N} \setminus \mathcal{O}_\omega} H_i \bar{g}_i$, the post-contingency system kinetic energy at nominal frequency,⁷ and,
- $\underline{E}_\omega = \left(\frac{f_{\min}}{f_0} \right)^2 E_\omega$, the system kinetic energy at the low frequency limit f_{\min} .

The quantity $\Delta e_\omega(x, t)$ can be estimated by the energy imbalance between load and generation up to time t due to the disturbances. Under a contingency scenario ω , let $\mathcal{O}_\omega = \{i | \tilde{a}_{i\omega} = 0\}$ be the set of buses with generation

⁷We assume that the machine MVA rating is equal to the maximum power output \bar{g}_i . We can also use estimates of system inertia constant using historical data [30], which includes inertia of load.

outage and let $\mathcal{A}_\omega = \mathcal{N}^G \setminus \mathcal{O}_\omega$ be the set of buses with available generators (in normal operation) having governors. Suppose that disturbance occurs at $t = 0$. According to our assumption on ramping of available generators, the subsequent increased mechanical power output from generator at bus $i \in \mathcal{A}_\omega$ is:

$$\Delta p_{i\omega}(x_i, t) = \begin{cases} \rho_i t, & \text{if } t < t_i \\ \bar{g}_i - x_i, & \text{if } t \geq t_i \end{cases}, \quad (4.5)$$

where ρ_i is the ramp rate limit and

$$t_i = \frac{\bar{g}_i - x_i}{\rho_i}, \quad (4.6)$$

is the time when generator i reaches its maximum output limit \bar{g}_i . For the convenience of analysis, we assume that total generation mechanical output will never exceed the load.⁸ Now, assume that total load is constant during the FSP and that pre-contingency dispatch x satisfies power balance constraint (4.1b). The increased mechanical power output from all available generators at time t would be a function of (x, t) :

$$\Delta p_\omega(x, t) = \begin{cases} \sum_{i \in \mathcal{A}_\omega} \Delta p_{i\omega}(x_i, t), & \text{if } t \leq t^* \\ \sum_{i \in \mathcal{O}_\omega} x_i, & \text{if } t > t^* \end{cases}, \quad (4.7)$$

where t^* is either

- the time when mechanical power output matches the load, if this occurs within the the duration of the FSP t^s , or,

⁸Although, in practice, generation mechanical output will exceed the load to restore frequency after the frequency nadir, this assumption has no impact on the estimation of amount of unserved energy in the adopted UFLS model.

- $t^* = t^S$, if mechanical power output fails to meet the load within t^S .

That is,

$$t^* = \min \left\{ t^S, \min_t \left\{ t \left| \sum_{i \in \mathcal{A}_\omega} \Delta p_{i\omega}(x_i, t) = \sum_{i \in \mathcal{O}_\omega} x_i \right. \right\} \right\}. \quad (4.8)$$

Given the mechanical power output function Δp_ω , t^* can be easily computed.

Note that without any UFLS, the system frequency reaches its nadir at t^* .

The change in stored kinetic energy at time t is the integral of power mismatch between generation and load up to time t , that is,

$$\Delta e_\omega(x, t) = \int_{\tau=0}^{\tau=t} \left(\Delta p_\omega(x, \tau) - \sum_{i \in \mathcal{O}_\omega} x_i \right) d\tau. \quad (4.9)$$

Let $\mathcal{B}_\omega(t) \subseteq \mathcal{A}_\omega$ be the set of generators having binding capacity constraints before t , that is,

$$\mathcal{B}_\omega(t) = \{i | t_i \leq t, i \in \mathcal{A}_\omega\},$$

and let $\mathcal{B}_\omega^c(t) = \mathcal{A}_\omega \setminus \mathcal{B}_\omega(t)$ be the complement set of $\mathcal{B}_\omega(t)$. Combine (4.5)

with (4.9), we obtain:

$$\begin{aligned} \Delta e_\omega(x, t) &= \sum_{i \in \mathcal{B}_\omega(t)} \left(\int_{\tau=0}^{\tau=t_i} \rho_i \tau d\tau + \int_{\tau=t_i}^{\tau=t} (\bar{g}_i - x_i) d\tau \right) \\ &\quad + \sum_{i \in \mathcal{B}_\omega^c(t)} \int_{\tau=0}^{\tau=t} \rho_i \tau d\tau - \int_{\tau=0}^{\tau=t} \sum_{i \in \mathcal{O}_\omega} x_i d\tau \\ &= \sum_{i \in \mathcal{B}_\omega(t)} \left(\frac{1}{2} \rho_i t_i^2 + (\bar{g}_i - x_i)(t - t_i) \right) \\ &\quad + \sum_{i \in \mathcal{B}_\omega^c(t)} \frac{1}{2} \rho_i t^2 - \sum_{i \in \mathcal{O}_\omega} x_i t. \end{aligned} \quad (4.10)$$

Note that Δe_ω is negative and non-increasing in t . Also, it reaches the minimum Δe_ω^* at time t^* , that is,

$$\Delta e_\omega^*(x) = \Delta e_\omega(x, t^*). \quad (4.11)$$

Therefore, enforcing minimum frequency constraints at time t^* ensures that (4.4) is satisfied, that is,

$$E_\omega + \Delta e_\omega^*(x) \geq \underline{E}_\omega. \quad (4.12)$$

4.3.2 Estimating the amount of UFLS

In the case that governor responses alone cannot suffice to satisfy the frequency constraint (4.12), UFLS needs to be deployed to maintain system frequency above the threshold f_{\min} . One approach to approximate the amount of UFLS [57] uses the shortfall in stored kinetic energy to meet the frequency threshold, that is,

$$e_\omega^s(x) = \max \{ \underline{E}_\omega - (E_\omega + \Delta e_\omega^*(x)), 0 \}. \quad (4.13)$$

Under this scheme, load shedding ends immediately when the mechanical power output meets the load. However, the load shed will remain disconnected from several minutes to several hours in practice. Therefore, we must estimate the amount of power shed that maintains the system frequency above the limit, and model the duration of the power shed.

In practice, there may be multiple frequency thresholds triggering different sets of under-frequency relays. There are also types of UFLS that are

triggered when the frequency remains below a certain threshold for a specified time interval, or triggered when the rate of change in frequency is sufficiently large. We therefore define the “theoretically optimal UFLS,” $p_\omega^s(x)$, to approximate the actual UFLS. Note that we assume a single frequency threshold f_{\min} . This is a widely used assumption for estimating under-frequency load shedding [57]. We assume that UFLS is triggered at \hat{t} , the time when the system frequency reaches the low frequency limit. That is,

$$\hat{t} = \min_t \{t | E_\omega + \Delta e_\omega(x, t) = \underline{E}_\omega\}. \quad (4.14)$$

The theoretically optimal UFLS is defined as the power imbalance between load and mechanical power output at \hat{t} :⁹

$$p_\omega^s(x) = \sum_{i \in \mathcal{O}_\omega} x_i - \Delta p_\omega(x, \hat{t}). \quad (4.15)$$

$p_\omega^s(x)$ represents the minimum amount of load shedding needed to ensure the frequency above the low frequency limit. Although p_ω^s can be computed efficiently given a dispatch x , the function $p_\omega^s(x)$ is non-convex. However, from numerical experiments, we find that the mapping between p_ω^s and e_ω^s is almost linear for scenarios having similar sizes of disturbance. Since $e_\omega^s(x)$ is convex (shown in Proposition 1 below), we use e_ω^s to approximate p_ω^s and ensure a convex approximation of the problem, that is,

$$\hat{p}_\omega^s(x) = \gamma_\omega e_\omega^s(x) + \sigma_\omega, \quad (4.16)$$

⁹Although we may achieve a lower amount of UFLS by disconnecting load earlier than \hat{t} , the theoretically optimal UFLS emulates the practice that a under-frequency relay is usually triggered with a certain frequency threshold.

where \hat{p}_ω^s is the approximated power shed. Non-negative coefficients γ_ω and σ_ω can be either obtained by *ex ante* simulation, or evaluated during the process of solving SED, as will be discussed in Section 4.4.

We use a single generator model to explain why (4.16) forms a reasonable approximation to $p_\omega^s(x)$. Assume that this generator has a pre-contingency dispatch \bar{x} , a maximum output limit \bar{g} , and a maximum ramp rate $\bar{\rho}$. Let the size of disturbance be L_ω , where $L_\omega = \sum_{i \in \Theta_\omega} x_i$. We assume the generator's headroom is greater than the size of disturbance, that is, $\bar{g} - \bar{x} > L_\omega$. We also assume $\frac{L_\omega}{\bar{\rho}} < t^s$; therefore, $t^* = \frac{L_\omega}{\bar{\rho}}$. We also suppose that UFLS is triggered. By (4.10), we have:

$$\Delta e_\omega(t) = \frac{\bar{\rho}t^2}{2} - L_\omega t. \quad (4.17)$$

Also, by (4.11), we obtain:

$$\Delta e_\omega^* = -\frac{L_\omega^2}{2\bar{\rho}}. \quad (4.18)$$

Substituting (4.18) into (4.13) gives:

$$e_\omega^s \approx \underline{E}_\omega - E_\omega + \frac{L_\omega^2}{2\bar{\rho}}. \quad (4.19)$$

Moreover, according to (4.15), p_ω^s is the power imbalance at time \hat{t} :

$$p_\omega^s = L_\omega - \bar{\rho}\hat{t}. \quad (4.20)$$

Substitute (4.17) into (4.14). Then \hat{t} satisfies the following equation:

$$E_\omega + \left(\frac{\bar{\rho}\hat{t}^2}{2} - L_\omega\hat{t}\right) = \underline{E}_\omega.$$

By solving the above equation, we obtain:¹⁰

$$\hat{t} \approx \frac{L_\omega - \sqrt{L_\omega^2 - 2\bar{\rho}(E_\omega - \underline{E}_\omega)}}{\bar{\rho}}. \quad (4.21)$$

Substituting (4.19) into the above equation gives:

$$\hat{t} \approx \frac{L_\omega - \sqrt{2\bar{\rho}e_\omega^s}}{\bar{\rho}}. \quad (4.22)$$

Combining with (4.20), we obtain:

$$p_\omega^s \approx \sqrt{2\bar{\rho}e_\omega^s}. \quad (4.23)$$

For a given scenario ω , if the associated kinetic energy shortfall $e_\omega^s(x)$ concentrates in a particular range, then (4.23) can be well-approximated by the linear function (4.16). Figure 4.4 in Section 4.5.2 provides a numerical example to support this observation.

Figure 4.2 illustrates the constructed UFLS model. The solid curve represents the mechanical power output of the generators, which is piecewise-linear. The breakpoint represents when a generation capacity constraint becomes binding. The frequency reaches the lower limit at \hat{t} , while the decrement in kinetic energy, represented by the area of solid shaded region, reaches the upper limit $E_\omega - \underline{E}_\omega$. The power imbalance at \hat{t} is the assumed power shed p_ω^s . In addition, the area with slanted lines represents e_ω^s , the kinetic energy shortfall to meet the frequency requirement f_{\min} without applying UFLS, which is used to approximate the amount of power shed p_ω^s .

¹⁰ \hat{t} is the smaller root since UFLS is triggered once the frequency reaches the low limit.

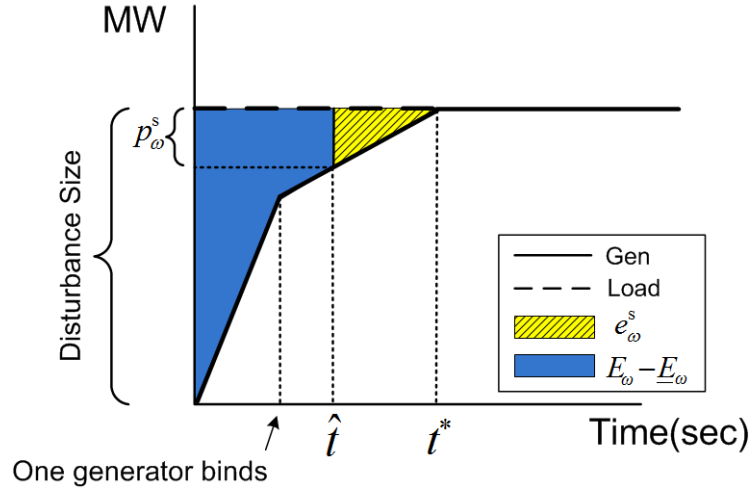


Figure 4.2: Illustration of constructed UFLS model. The solid curve represents the mechanical power output of the generators, which is piece-wise linear. The breakpoint represents when a generation capacity constraint becomes binding. The frequency reaches the lower limit at \hat{t} , while the decrement in kinetic energy, represented by the area of solid shaded region, reaches the upper limit $E_\omega - \underline{E}_\omega$. The power imbalance at \hat{t} is the theoretically optimal UFLS p_ω^s .

Proposition 1. $\hat{p}_\omega^s(x)$ is a convex function in x .

Proof. In (4.5), since $\rho_i t$ is non-decreasing, $\Delta p_{i\omega}(x_i, t)$ can be expressed as a pointwise minimum function:

$$\Delta p_{i\omega}(x_i, t) = \min\{\rho_i t, \bar{g}_i - x_i\}.$$

This function is concave in x_i . Therefore, $\sum_{i \in \mathcal{A}_\omega} \Delta p_{i\omega}(x_i, t)$ is also concave in x . By the same reasoning of pointwise minimum, $\Delta p_\omega(x, t)$ is also concave in x . By (4.7), we know that:

$$\Delta p_\omega(x, t) = \sum_{i \in \mathcal{O}_\omega} x_i, t > t^*.$$

Therefore, $\Delta e_\omega^*(x)$ can also be expressed as:

$$\Delta e_\omega^*(x) = \int_{t=0}^{t=t^s} \left(\Delta p_\omega(x, t) - \sum_{i \in \mathcal{O}_\omega} x_i \right) dt.$$

Since for any $t \in [0, t^s]$, $\Delta p_\omega(x, t)$ is a concave function in x , the integral over t of $\Delta p_\omega(x, t)$ is also concave. The maximum operator in (4.13) implies that $e_\omega^s(x)$ is a convex function in x . Since γ_ω in (4.16) is non-negative, $\hat{p}_\omega^s(x)$ is therefore convex. \square

Some governors do not respond to decline in system frequency immediately, but with some response delay. This initial response delay can be incorporated in our model. Given the initial delay t_i^d for governor of generator at bus i , the increased mechanical power output after the disturbance would

be:

$$\begin{aligned} \Delta p_{i\omega}(x_i, t) &= \begin{cases} 0, & \text{if } t < t_i^d \\ \rho_i(t - t_i^d), & \text{if } t_i^d \leq t < t_i \\ \bar{g}_i - x_i, & \text{if } t \geq t_i \end{cases} \\ &= \min \{ \max \{ 0, \rho_i(t - t_i^d) \}, \bar{g}_i - x_i \}. \end{aligned} \quad (4.24)$$

Thus $\Delta p_{i\omega}(x_i, t)$ is still a concave function in x_i and the convexity of $e_\omega^s(x)$ is preserved. Assume that $t^* > t_i^d$ for $i \in \mathcal{A}_\omega$. With a similar derivation to incorporate governor delay, the decrement in kinetic energy at time t^* is:

$$\begin{aligned} \Delta e_\omega^*(x) &= \sum_{i \in \mathcal{B}_\omega(t^*)} \left(\frac{1}{2} \rho_i (t_i - t_i^d)^2 + (\bar{g}_i - x_i)(t^* - t_i) \right) \\ &\quad + \sum_{i \in \mathcal{B}_\omega^c(t^*)} \frac{1}{2} \rho_i (t^* - t_i^d)^2 - \sum_{i \in \mathcal{O}_\omega} x_i t^*. \end{aligned} \quad (4.25)$$

As we discuss in Section 4.4, the L-shaped method (algorithm) is applied to solve SED. The method requires repeated evaluation of the subgradient of $\hat{p}_\omega^s(x)$. Appendix A develop an analytical expression for that subgradient based on one particular directional derivative. This analytical expression can significantly improve computational efficiency.

4.4 Decomposition Algorithm

Slyke and Wets [61] develop the L-shaped method to exploit the special structure of a stochastic linear program (See also Benders [7]). Birge and Tang generalize this method to the case of stochastic convex program [8]. Since $\hat{p}_\omega^s(x)$ is convex, the optimal value of the second-stage problem Q_ω is a convex functions of (x, r) [22]. Therefore, the L-shaped method can be applied

to solve SED. The main idea of the L-shaped method is to represent sub-problems Q_ω with their outer linearizations in (4.1), so that Q_ω can be solved independently for each scenario. The decomposition algorithm is described below as Algorithm 2.¹¹

At the beginning of the algorithm, we need to set the coefficients γ_ω and σ_ω in (4.16) to some initial values. A reasonable initial value for γ_ω is between 0 to 1, where the unit is sec^{-1} , since the value of p_ω^s is usually smaller than the value of e_ω^s according to numerical experiments. A good initial guess for σ_ω is 0, because there would be no load-shedding with e_ω^s of 0. During the first several iterations, in addition to evaluating e_ω^s , we compute p_ω^s using (4.15) to obtain γ_ω and σ_ω by linear regression. From numerical experiments, we observe that the correlation between e_ω^s and p_ω^s tends to be strong under similar sizes of contingency (see Figure 4.4); thus a moderate number of samples of e_ω^s and p_ω^s is likely to provide reasonable estimates of γ_ω and σ_ω . Note that we only update γ_ω and σ_ω once in the algorithm. Therefore, after the update, the algorithm becomes a standard decomposition algorithm for a non-linear program, for which the convergence is guaranteed [25] [32].

In the L-shaped method, optimality cuts (4.26b) should form lower bounds of Q_ω . However, updating the coefficients for scenario ω changes the function Q_ω . As a result, existing optimality cuts associated with ω might no

¹¹Note that we assume the second-stage problem (4.2) is feasible for every feasible first-stage decisions. This can be easily achieved by adding a artificial variable, with a large objective function penalty, to each constraint in (4.2g). Therefore no feasibility cuts are needed.

Algorithm 2 Solving SED Model

Input: Network data, optimality tolerance ϵ .

Output: ϵ -optimal (x^*, r^*) and the associated cost z^* .

$m \leftarrow 1, z^* = +\infty$.

loop

Solve the relaxed master program:

$$\min_{x, r, \theta} \pi_0 (C_e(x) + C_r(r)) + \sum_{\omega \in \Omega} \pi_\omega \theta_\omega \quad (4.26a)$$

s.t. Constraints (4.1b)-(4.1h)

$$\begin{aligned} \theta_\omega \geq & \sum_{i \in \mathcal{N}} \psi_\omega^k \beta_{i\omega}^k x_i + \sum_{i \in \mathcal{N}} \tilde{a}_{i\omega} ((\lambda_{i\omega}^k + \mu_{i\omega}^k)(x_i + r_i^{\text{SP}} \\ & + r_i^{\text{FS}}) + \lambda_{i\omega}^k r_i^{\text{NS}}) + \eta_\omega^k, \forall \omega \in \Omega, k = 1, \dots, m. \end{aligned} \quad (4.26b)$$

Let (\hat{x}, \hat{r}) be the optimal solution of the relaxed master program and let \underline{z} be the corresponding optimal value. For all $\omega \in \Omega$, compute $e_\omega^s(x)$ and $\hat{p}_\omega^s(x)$ using (4.13) and (4.16), respectively. Form the right-hand sides of (4.2) and solve $Q_\omega(\hat{x}, \hat{r})$. Then compute the associated cost \hat{z} using (4.1a).

if $\hat{z} \leq z^*$ **then**

$z^* \leftarrow \hat{z}$

$(x^*, r^*) \leftarrow (\hat{x}, \hat{r})$

end if

if $z^* - \underline{z} \leq \epsilon \min(z^*, \underline{z})$ **then**

Stop: (x^*, r^*) is the ϵ -optimal solution.

end if

For each $\omega \in \Omega$, compute $\beta_{i\omega}(\hat{x})$ using equation (A.6) in Appendix A. Augment the set of optimality cuts in (4.26b) with cut coefficients:

$$\beta_{i\omega}^m = \beta_{i\omega}(\hat{x}), \psi_\omega^m = \hat{\psi}_\omega, \lambda_{i\omega}^m = \hat{\lambda}_{i\omega}, \mu_{i\omega}^m = \hat{\mu}_{i\omega},$$

where $\hat{\psi}_\omega, \hat{\lambda}_{i\omega}$, and $\hat{\mu}_{i\omega}$ are the optimal dual variables of constraints (4.2b), (4.2e), and (4.2h), respectively. The cut intercept η_ω^m is:

$$\begin{aligned} \eta_\omega^m = & Q_\omega(\hat{x}, \hat{r}) - \sum_{i \in \mathcal{N}} \psi_\omega^m \beta_{i\omega}^m \hat{x}_i - \sum_{i \in \mathcal{N}} \tilde{a}_{i\omega} ((\lambda_{i\omega}^m + \mu_{i\omega}^m) \\ & (\hat{x}_i + \hat{r}_i^{\text{SP}} + \hat{r}_i^{\text{FS}}) + \lambda_{i\omega}^m \hat{r}_i^{\text{NS}})). \end{aligned}$$

$m \leftarrow m + 1$

end loop

longer be lower bounds of Q_ω and thus should be removed. However, under certain conditions, as described in Appendix B, the existing optimality cuts can be preserved. Some computational effort can be saved by keeping the existing cuts since they still provide useful information of the updated second-stage problem. In the context of solving successive SEDs for consecutive dispatch intervals, the updated coefficients can serve as good initial values for solving the SED for the subsequent dispatch interval. To simplify the presentation of the L-shaped method, we do not include this procedure of updating γ_ω and σ_ω in Algorithm 2.

Since the cost of involuntary load shedding is dominant in (4.2a) [29], we may ignore the difference in pre- and post-contingency energy-production costs. That is, the approximated objective function of the re-dispatch model (4.2) is $ct^L(l_\omega^S + \sum_{i \in \mathcal{N}} l_{i\omega}^R)$, and the second-stage energy costs are aggregated with the first-stage energy costs. This approximation simplifies the second-stage problems (4.2) from quadratic programs into linear programs, which significantly reduces the computational burden for large problems.

4.5 Numerical Examples

4.5.1 Validation of simplified frequency model

To validate the simplified frequency model developed in Section 4.3, we perform a case study by comparing our model with the transient simulation results produced by PowerWorld. A nine-bus system with four generators is used. Given a dispatch x , the change in stored kinetic energy is computed

Table 4.1: Comparison of simulated and estimated frequency nadirs under different sizes of disturbances (frequency units are Hz.)

x_1 (MW)	PowerWorld	Estimation	error (%)
30	59.206	59.200	0.01
35	58.894	58.917	0.04
40	58.538	58.567	0.05
45	58.132	58.155	0.04
50	57.674	57.679	0.01

using (4.10). Combining with (4.3), the estimated frequency at time t would be:

$$f(t) = \frac{1}{2\pi} \sqrt{\frac{2(E_\omega + \Delta e_\omega(x, t))}{I}}.$$

We consider generator outage at bus 1. The pre-contingency dispatch is $x_2 = 140\text{MW}$, $x_3 = 100\text{MW}$, $x_4 = 60\text{MW}$. Different values of pre-contingency dispatch x_1 , for generator at bus 1, are used to simulate the effect of various magnitudes of disturbances. Suppose that $\bar{g}_i = 150\text{MW}$, $H_i = 5\text{s}$, $t_i^d = 0.2\text{s}$, and $\rho_i = 168\text{MW}/\text{min}$, $\forall i \in \mathcal{N}^G$.¹² Table 4.1 compares the results, which indicate that the proposed frequency model does accurately estimate frequency nadirs. The errors are all under 0.05%. For a disturbance size of 35MW, around 12 % of the load, Figure 4.3 plots the estimated and simulated frequency, indicating the proposed frequency model also accurately evaluate system frequency in the time domain.

¹²We use governor model “GGOV1” in the version 15 of PowerWorld. All governor parameters are set to default except for Ropen = 0.015, Kturb = 1.25, and Trate = 150MW.

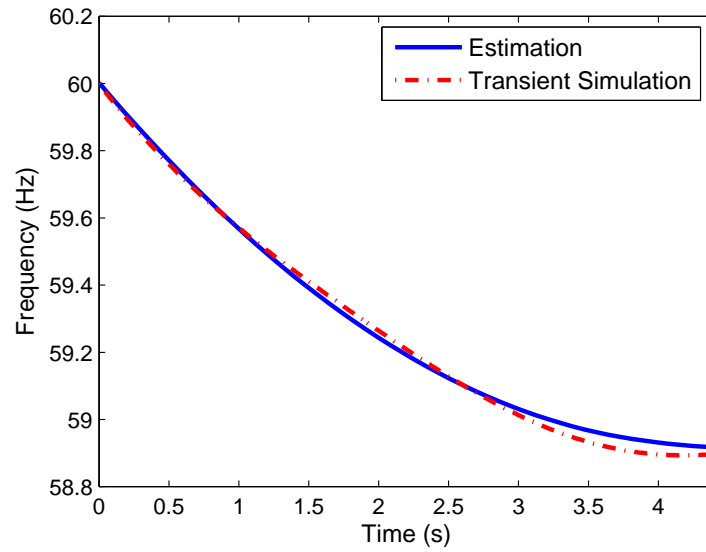


Figure 4.3: Comparison of estimated and simulated frequency in time domain. This figure shows that the simplified frequency model developed in Section 4.3 does accurately estimate the frequency, compared to transient simulation results using PowerWorld.

4.5.2 IEEE 118-bus Reliability Test System

Power flow data for the IEEE 118-bus Reliability Test System together with pre-defined contingency scenarios are used here to compare SED and conventional economic dispatch. Each scenario corresponds to a single generator outage. Suppose that the system inertia constant is 6s. The outage probabilities for the 46 in-service generators, estimated from the reliability data in [54], range from 1×10^{-4} to 1.1×10^{-3} during the operational time frame of 30 minutes. Other important problem specifications are:

- The minimum frequency is set as 57.5Hz, with a nominal frequency of 60Hz.
- The total load in the system is 3668 MW, and on-line generation capacity is 4747 MW.
- The penalty c is set as \$10,000/MWh with load-shedding duration t^L of one half hour.

We use generic generation cost data provided in [54]. This example focuses on the benefits of modeling frequency control and geographical allocation of spinning reserves, and thus omits fast-start or non-spinning reserves.¹³ Besides 194 base-case transmission constraints, we add 10 binding and near-binding $N - 1$ transmission security constraints by solving security-constrained

¹³We therefore assume \bar{g}_i is 0 for any generator with $u_i = 0$. Also, reserve restoring constraint (4.2g) is therefore ignored.

optimal power flow (SCOPF) in PowerWorld. We use Gurobi 4.5 as the linear and quadratic program solver, together with the Python language to implement the decomposition.

In addition to solving SED (4.1), the economic dispatch (ED) model with fixed reserve requirements is solved for comparison [5]. The spinning reserve requirement of the ED model is 400 MW, which is the capacity of the largest unit. We then evaluate the post-contingency operation cost of the ED solution by solving the second-stage problems (4.2). Using a laptop computer on Windows 7 with 2.0 GHz dual core CPU and 3GB RAM, the CPU time for solving SED is 24.19 seconds. It takes 56 outer iterations to converge to an optimality tolerance of 0.01%. Table 4.2 shows that the saving of expected cost using the stochastic solution is 3.79%. Note that there are only three units producing at their maximum capacity in the stochastic solution, compared to 21 units in the deterministic solution. In other words, in the stochastic solution, more units are available for responding to the disturbance simultaneously and, consequently, the expected UFLS is about 45% lower, compared to the deterministic solution. Besides better frequency response, SED improves the geographical allocation of reserves by incorporating post-contingency transmission constraints, which contribute to in 20% less expected load shedding deployed in the RDP. The trade-off for better security is a 4.72% higher pre-contingency cost for the stochastic solution.

Figure 4.4 shows the scatter plot of p_{ω}^s and e_{ω}^s , obtained at the first two outer iterations. We collect contingencies with similar sizes into a group,

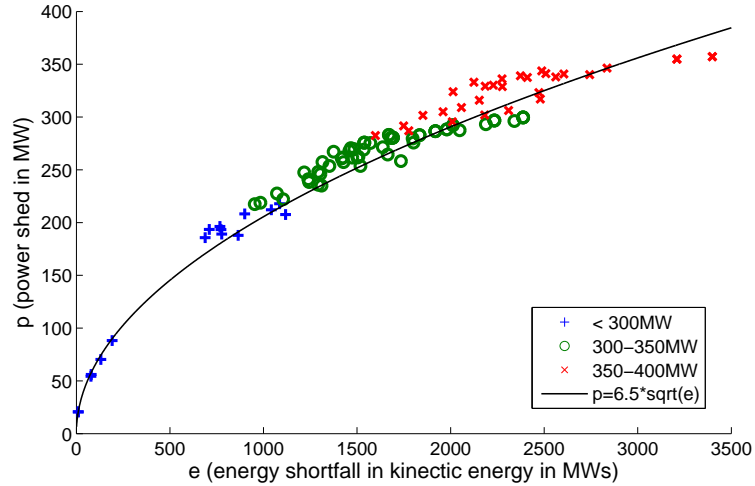


Figure 4.4: e_{ω}^s versus p_{ω}^s for the IEEE 118-bus Reliability Test System. This figure indicates that p_{ω}^s is approximately linear in e_{ω}^s for each group of contingencies. This finding suggests that we can reasonably approximate p_{ω}^s using e_{ω}^s for ensuring the convexity of SED, as described by the (4.16) in Section 4.3.

and each type of marker in the figure represents such a group. Interestingly, the figure shows that p_{ω}^s is roughly proportional to $\sqrt{e_{\omega}^s}$, which is consistent with (4.23). This figure also indicates that p_{ω}^s is approximately linear in e_{ω}^s for each group of contingencies. This finding suggests that we can reasonably approximate p_{ω}^s using e_{ω}^s for ensuring the convexity of SED, as described by the (4.16) in Section 4.3.

Table 4.2: Results for the 118 bus system

	SED	ED
Expected cost (\$/h)	57584.65	59852.86
EUFLS ¹	0.60	1.35
EUL in RDP ²	0.73	0.91
Pre-Contingency cost (\$/h)	50993.22	48695.70
Spin. Reserve Procured (MW)	449.56	400.00

¹ Expected UFLS (MW) triggered in the FSP.

² Expected load shedding (MW) deployed in the RDP.

4.5.3 ERCOT system

The ERCOT network and load data for summer 2011 together with generic costs¹⁴ and pre-defined contingency scenarios are used to compare SED and conventional economic dispatch. Assume the system inertia constant is 4s [36]. The probability of a single generation contingency during the operational time frame of one half hour, estimated by the generating availability data system (GADS) provided by NERC [46], ranges from 8.8×10^{-5} to 6.8×10^{-3} , depending on the type of generator. These probabilities are then used to construct contingency scenarios, as discussed in the next paragraph. The minimum frequency is set as 59.5 Hz¹⁵, with the nominal frequency of 60 Hz. The total load in the system is 59.60 GW, on-line generation capacity (without wind capacity) is 59.68 GW, and the forecasted wind output is 4431 MW. The penalties c are set \$30,000/MWh with load-shedding duration

¹⁴For instance, for a 600MW coal-fired generation unit, the marginal cost curve, represented as a linear function, ranges from \$24-30/MWh. As another example, for a 400MW gas-fired steam turbine, the marginal cost curve, also represented as a linear function, ranges from \$63-71/MWh.

¹⁵The highest frequency thresholds in North America range from 59.3 to 59.7 Hz [36].

t^L of one half hour.¹⁶ In this example, we focus on the benefits of modeling frequency control and geographical allocation of spinning reserves and thus do not include fast-start and non-spinning reserves. Also, besides 7186 base-case transmission constraints, we add 10 binding and near-binding $N - 1$ transmission security constraints, obtained from solving security-constrained optimal power flow (SCOPF) in PowerWorld with the given power flow and generic cost data.

We use historical wind data in the ERCOT during 2008 and 2009 to estimate the probabilities and sizes of significant wind ramp events in the operational time frame of an half hour. (Note that in real-time operation, we can utilize a wind ramp forecasting tool to obtain more accurate estimates of the likelihood of large wind events [2].) Given a sequence of wind power output data sampled with fixed time interval $\{p_k\}$, $k = 1, \dots, n$, and assume m samples are taken in 30 minutes. Define set $B = \{k | \underline{p} \leq p_k < \bar{p}\}$ and the set $A = \{k | \underline{\delta} \leq p_k - p_{k+m} < \bar{\delta}\}$. The conditional probability of a wind ramp event is estimated as

$$P(\underline{\delta} \leq \text{ramp size} < \bar{\delta} | \underline{p} \leq \text{wind power output} < \bar{p}) \approx \frac{|A|}{|B|},$$

where the size of the ramp is represented by the mean of samples in A . Table 4.3 summarizes the derived wind ramp sizes and probabilities conditioned on the forecasted wind output. Note that the decreased wind output is dis-

¹⁶The parameter c roughly corresponds to the reliability standard “one day in ten years” [28].

Table 4.3: Estimated conditional probability of significant wind ramp event

Ramp Range (MW)	Avg. Ramp Size (MW)	Prob.(%)
> 2500	2278.64	0.014
1500-2000	1634.27	0.047
1000-1500	1132.34	0.038
500-1000	651.83	4.48

tributed to each wind farm, which is treated as a generator, proportional to its forecasted output. There are 87 on-line wind farms in the power flow data.

Each contingency scenario consists of generation outages or a realization of significant wind ramp, or both. We assume that wind events and generation outages are independent. The largest generation contingency is the loss of two co-located nuclear units, totaling 2443 MW of capacity.¹⁷ To make the number of scenarios manageable, only generation contingencies greater than 900 MW in capacity are considered. Furthermore, scenarios with co-occurrence of large wind ramps and generation outages are only incorporated if the associated probabilities are greater than 10^{-6} . There are a total of 36 contingency scenarios after applying this scenario reduction rule.

4.5.3.1 Impact of various risk of wind ramps

The probabilities of wind ramps in Table 4.3 are scaled to increased levels to simulate the system under different risk of significant wind events. The

¹⁷Although the size of contingency is determined by the actual production of generators, we use capacities to approximate the actual generation outputs. This approximation is reasonable particularly for units with large capacity, since these units usually produce at their maximum output limit due to their low operation costs.

first column of Table 4.4 shows the scaled probability of the largest ramp. All other probabilities are scaled proportionally. The ED model with a spinning reserve requirement of 2443 MW (size of the largest generation contingency) is solved for comparison. Table 4.4 shows that SED produces about 1.5% of cost saving over the ED for each case. The fourth and fifth columns of Table 4.4 show that the expected UFLS is significantly reduced using the SED solution, and suggests that the cost savings are mainly due to the improved system frequency response. Apparently, SED procures more spinning reserve under higher probabilities of wind ramps to reduce risk associated with involuntary load shedding. However, in all cases, the procured amounts are lower than the spinning reserve requirement in the ED model. Interestingly, in some cases, a lower value of spinning reserves results in less load shedding in the RDP, probably because of a better geographical allocation.

4.5.3.2 Impact of increased level of wind integration

We simulate the impact of expansion in installed wind capacity by increasing the size of wind ramps by 25%: the largest wind ramp is now 2848 MW. This corresponds to the estimate reported in [3], which predicts that, in the future, the largest wind ramp-down will exceed 2800 MW with 15,000 MW of installed wind capacity in the ERCOT region. Again, we scale the probabilities of wind ramps in Table 4.3 to simulate the system under different risk of significant wind events. Two ED models are solved: one with spinning reserve requirement of 2443 MW (size of the largest generation

Table 4.4: Impact of various risk of wind ramps

Prob. ¹	Exp. Cost ²		EUFLS ³		EUL in RDP ⁴		SP ⁵
	SED	Saving ⁶	SED	ED	SED	ED	
0.014	1853644.8	1.63	0.802	2.644	0.244	0.004	1634.3
0.056	1857092.4	1.46	0.790	2.646	0.008	0.018	2278.6
0.112	1857465.0	1.45	0.768	2.644	0.026	0.040	2278.6
0.168	1857355.6	1.46	0.760	2.642	0.038	0.060	2278.6
0.224	1857283.4	1.48	0.786	2.638	0.050	0.080	2278.6
0.280	1857704.4	1.46	0.788	2.636	0.062	0.100	2278.6

¹ Probability (%) of the largest wind ramp.

² Expected cost in \$/h.

³ Expected UFLS (MW) triggered in the FSP.

⁴ Expected load shedding (MW) deployed in the RDP.

⁵ Spinning reserve (MW) procured by the SED solution.

⁶ Cost saving in % compared to ED.

contingency), and the other with reserve requirement of 2848 MW (size of the largest wind ramp). Figure 4.5 shows that the cost savings using stochastic approach increase with higher risk of wind events. The cost savings range from 1.49% to 3.17%, compared to the ED model with 2443 MW of reserves. On the other hand, the cost savings range from 1.29% to 2.08%, compared to the ED model with 2848 MW of reserves. Interestingly, the SED produces a much more secure dispatch under high risk of wind ramps in terms of expected load shedding, as shown in Figure 4.6. Note that the total amount of spinning reserves of the SED solution is close to 2848 MW when the probability of the largest wind ramp exceeds 5×10^{-4} , as shown in Figure 4.7. However, improved frequency response together with better allocation of reserves of the

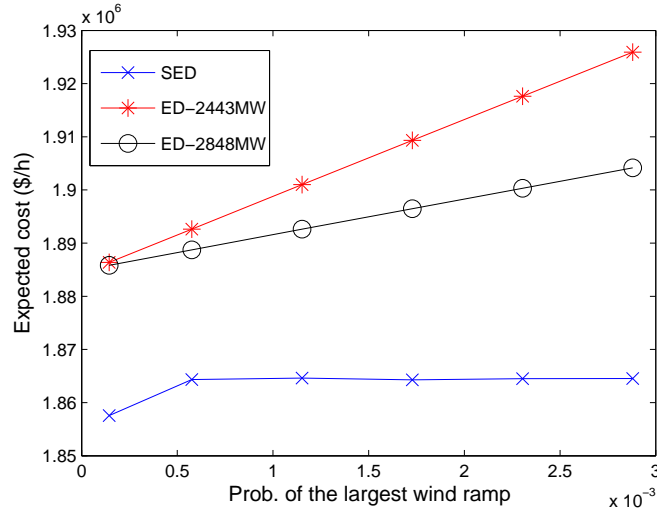


Figure 4.5: Expected cost in the ERCOT example. This figure shows that the cost savings using stochastic approach increase with higher risk of wind events. This finding implies that as stochastic variability increases for supply, deterministic solutions compare more and more unfavorably with stochastic solutions.

stochastic solutions lower the expected post-contingency operation costs. This numerical experiment implies that the stochastic approach is more useful as the penetration levels of wind increase. As stochastic variability increases for supply, deterministic solutions compare more and more unfavorably with stochastic solutions.

4.5.3.3 Computational costs

With an optimality tolerance of 0.05%, the average CPU time for solving SED is 278.8 seconds, with 27 outer iterations on average. Note that the reported CPU time includes the time for loading network data and creating optimization models, which takes 132 seconds on average. In the case of solving

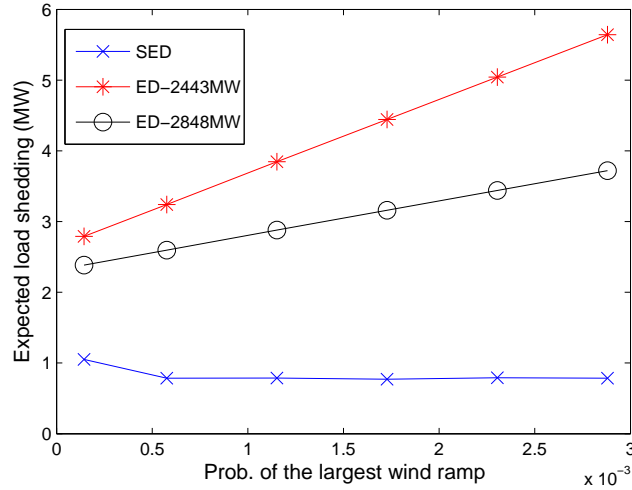


Figure 4.6: Expected load shedding in the ERCOT example. This figure shows that the SED produces a much more secure dispatch under high risk of wind ramps in terms of expected load shedding.

the model repeatedly with different system conditions, the created optimization models can be re-used after modification. Also, since the subproblem for each scenario (4.2) can be solved independently, parallel computing could be used for applying SED in real-time operation. For example, if subproblems are distributed to five computing resources, we may be able to solve this problem instance within one minute.

4.6 Conclusion

This chapter has proposed a new model, the stochastic economic dispatch model (SED), to address issues regarding reserve allocation in power system operations, namely, variability of wind, deliverability of reserves, and

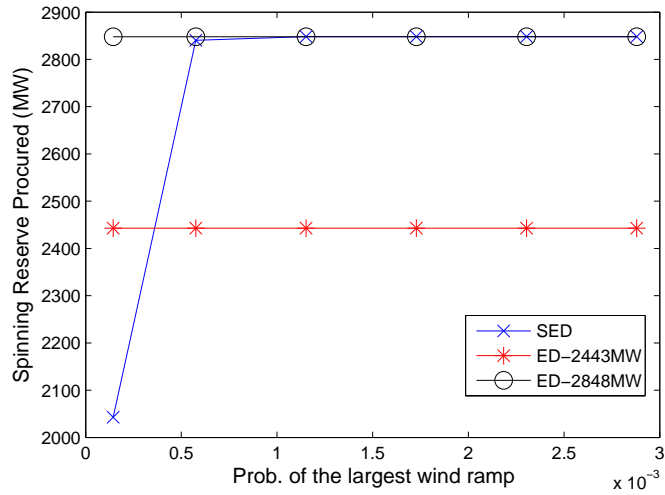


Figure 4.7: Spinning reserves procured in the ERCOT example. The total amount of spinning reserves of the SED solution is close to 2848 MW when the probability of the largest wind ramp exceeds 5×10^{-4} . However, improved frequency response together with better allocation of reserves of the stochastic solutions lower the expected post-contingency operation costs.

adequacy levels of primary frequency control. These issues are particularly important in the face of increasing levels of wind power penetration. SED, formulated as a two-stage stochastic convex program, searches for the optimal energy dispatch and reserve allocation considering important operational constraints. In particular, we propose a convex function to estimate under-frequency load shedding by incorporating frequency dynamics. The proposed model solves efficiently through decomposition. By comparing to transient simulation using PowerWorld, we also show that the convex formulation of system dynamics produces reasonable estimates on the average frequency. Numerical results imply that SED generates energy and reserves schedules that significantly reduce expected operation costs, compared to the current approaches of allocating reserves in most electricity markets. The results also suggest that the cost savings using SED becomes more significant under higher levels of wind integration.

Chapter 5

Conclusion

5.1 Summary

This dissertation has developed new methods for improving the efficiency of electricity markets. Specifically, we develop new approaches for market monitoring and construct a new economic dispatch model for efficiently allocating reserve.

Market monitoring is a critical and challenging task for keeping electricity markets well-functioning. Market monitors usually adopt a variety of indices to assess the levels of competition. These indices typically quantify market power, defined as a market participant’s ability to profitably shift prices away from competitive levels. However, most currently used indices have only weak connections to the economic incentives of exercising market power. This is particularly the case for indices that incorporate the effects of transmission constraints. As a result, these *ad hoc* indices might produce unreliable or even misleading market monitoring results.

To overcome the shortcomings of existing indices, Chapter 2 proposes four new market power indices based on “small-signal” analysis:

- the transmission-constrained inverse residual demand derivative,

- the derivative of price with respect to injection,
- the estimated price-cost mark-up, and
- the estimated wealth transfer over competitive levels.

Unlike the *ad hoc* indices adopted in most markets, these indices are built on economic principles and a faithful representation of transmission constraints. These indices are based on the assumption that each market participant owns all its assets at the same location in a transmission network. Although this assumption is restrictive, the theoretical and computational development of these indices provide a basis for the general case described in Chapter 3: a market participant owns assets at different locations. Because of the complexity of strategic behavior of market participants, market monitors usually need different approaches to analyze market power. Therefore, in addition to the small-signal analysis, Chapter 3 also presents a “large-signal” analysis to help market monitors interpret the market outcomes from a different perspective. To summarize, the market power indices proposed in Chapter 3 are

- the transmission-constrained market power index (TCMPI), which estimates the average price mark-ups above competitive levels for a given firm, and,
- the transmission-constrained residual supply index (TCRSI), which assesses the extent to which a firm’s supply is necessary to meet demand under transmission constraints.

These two indices provide different insights about market power: the TCMPI *detects* the *exercise* of market power, and the TCRSI *predicts* the *possession* of market power. Techniques for efficiently computing these indices are also presented. Numerical results show that the TCMPI indeed produces reasonable estimates on price mark-ups above competitive levels. In addition, numerical examples indicate that the TCRSI is a more accurate index in transmission-constrained markets, compared to the RSI. From both qualitative and quantitative analyses, we conclude that both these indices should be incorporated into the market power analysis, because their different insights about market power can help market monitors better judge the competitive conditions in the market.

Another important issue we address is the reserve allocation problem in both system and market operations. Most approaches adopted in current markets for allocating reserves lacks of representation of key issues such as variability of wind, deliverability of reserves, and sufficiency of primary frequency control. Therefore, current approaches may lead to an improper allocation of reserves that may cause unnecessary operation costs or threat power system reliability. To enhance the allocation of reserves, we propose the stochastic economic dispatch model (SED) in Chapter 4. Unlike other stochastic dispatch models in the literature, which use the post-contingency “steady-state” equilibrium to characterize the post-contingency operation, the proposed model describes the dynamic process of disturbance recovery. Consequently, the reliability benefits of various types of reserves are represented more faithfully in

the optimization model. To summarize, our main contributions are as follows:

- We propose the idea of incorporating frequency dynamics into a stochastic optimization framework, so that primary frequency control and geographical allocation of reserves can be co-optimized to minimize the expected operation cost.
- We develop a new convex formulation for estimating under-frequency load shedding based on principles of frequency dynamics.
- We apply a decomposition algorithm to efficiently solve the proposed model for large-scale systems.

The numerical examples based on medium- and large-scale systems show the value of the proposed formulation. That is, SED can indeed reduce the expected operation costs, compared with the conventional economic dispatch. Numerical results suggest that the cost savings range from 1-3%, which are significant savings for a daily operation. In most cases, the cost savings are mainly due to the improved reserve allocation that reduces the risks of involuntary load shedding. The results also suggest that the cost savings using SED becomes more significant with higher levels of wind power penetration, whereas the conventional economic dispatch is no longer sufficient for reliable and economical system operations.

5.2 Future Research

In Chapter 2 and 3, we mainly focus on the theoretical construction of market power indices. Although several numerical experiments are provided, the offer data of market participants are based on either marginal costs or theoretical profit-maximizing offers. We believe that applying the new indices on real market data would be an interesting and important research topic. In particular, we should pay attention to those market-clearing intervals with exceptionally high prices and then examine whether the proposed indices are able to distinguish the high competitive prices from the exercise of market power. Note that forward contract should be considered to properly model the incentives of profit maximization. Also, since the indices serve as estimates of market power, statistical analysis using historical market data might be useful for quantifying how reliable these indices are. This analysis can help the market monitor interpret the indices with higher confidence.

For the stochastic economic dispatch model, there are several issues that have not been addressed in this dissertation. For example, we use a simple rule to generate contingency scenarios, which selects scenarios based on contingency sizes and probabilities. However, there may exist some contingencies that are small in magnitude yet cause a significant amount of involuntary load shedding due to unresolvable transmission congestion. On the other hand, some contingencies with relatively small probabilities might cause a large amount of load shedding and thus should be included in the optimization model. In principle, we should choose scenarios based on their probability-weighted costs. The de-

velopment of such a method would be an important and interesting research topic for enhancing the efficiency of SED.

As another example, in the numerical examples, we include only a fraction of transmission security constraints. To incorporate all the $N - 1$ transmission security constraints in both pre- and post-contingency problems, it is necessary to couple SED with a contingency analysis module that produces transmission constraints on the fly. However, this combination may greatly increase the computational effort. One possible solution is to develop constraint pruning techniques to reduce the number of security constraints. In addition, parallel computing can be utilized to alleviate the computational burden, since SED can be decomposed to independent sub-problems through the L-shaped method.

Appendices

Appendix A

Derivation of Subgradient of \hat{p}_ω^S

We start with deriving a subgradient of Δe_ω^* , which is the key part of the whole derivation. By (4.10) and (4.11), we have:

$$\begin{aligned} \Delta e_\omega^*(x) &= \sum_{i \in \mathcal{B}_\omega(t^*)} \left(\frac{1}{2} \rho_i t_i^2 + (\bar{g}_i - x_i)(t^* - t_i) \right) \\ &\quad + \sum_{i \in \mathcal{B}_\omega^c(t^*)} \frac{1}{2} \rho_i t^{*2} - \sum_{i \in \mathcal{O}_\omega} x_i t^*. \end{aligned}$$

Assume that sets $\mathcal{B}_\omega(t^*)$ and $\mathcal{B}_\omega^c(t^*)$ remain unchanged under a small perturbation of x . By partially differentiating $\Delta e_\omega^*(x)$ with respect to $x_s, s \in \mathcal{B}_\omega(t^*)$, we obtain:

$$\begin{aligned} \frac{\partial \Delta e_\omega^*}{\partial x_s}(x) &= \frac{\partial t_s}{\partial x_s} (\rho_s t_s - (\bar{g}_s - x_s)) - (t^* - t_s) \\ &\quad + \frac{\partial t^*}{\partial x_s} \left(\sum_{i \in \mathcal{B}_\omega(t^*)} (\bar{g}_i - x_i) + \sum_{i \in \mathcal{B}_\omega^c(t^*)} \rho_i t^* - \sum_{i \in \mathcal{O}_\omega} x_i \right) \end{aligned} \tag{A.1}$$

By (4.7) and (4.8), we know that:

$$\sum_{i \in \mathcal{A}_\omega} \Delta p_{i\omega}(x, t^*) = \sum_{i \in \mathcal{B}_\omega(t^*)} (\bar{g}_i - x_i) + \sum_{i \in \mathcal{B}_\omega^c(t^*)} \rho_i t^* = \sum_{i \in \mathcal{O}_\omega} x_i,$$

under the case that mechanical power output meets the load. Otherwise, $t^* = t^s$ according to (4.8), which implies $\frac{\partial t^*}{\partial x_s} = 0$. Therefore,

$$\frac{\partial t^*}{\partial x_s} \left(\sum_{i \in \mathcal{B}_\omega(t^*)} (\bar{g}_i - x_i) + \sum_{i \in \mathcal{B}_\omega^c(t^*)} \rho_i t^* - \sum_{i \in \mathcal{O}_\omega} x_i \right) = 0.$$

Together with (4.6), equation (A.1) is reduced to:

$$\frac{\partial \Delta e_\omega^*}{\partial x_s}(x) = t_s - t^*. \quad (\text{A.2})$$

Similarly, $\frac{\partial \Delta e_\omega^*}{\partial x_s}(x)$ for $s \in \mathcal{B}_\omega^c(t^*)$ is:

$$\begin{aligned} \frac{\partial \Delta e_\omega^*}{\partial x_s}(x) &= \frac{\partial t^*}{\partial x_s} \left(\sum_{i \in \mathcal{B}_\omega(t^*)} (\bar{g}_i - x_i) + \sum_{i \in \mathcal{B}_\omega^c(t^*)} \rho_i t^* - \sum_{i \in \mathcal{O}_\omega} x_i \right) \\ &= 0, \end{aligned} \quad (\text{A.3})$$

and $\frac{\partial \Delta e_\omega^*}{\partial x_s}(x)$ for $s \in \mathcal{O}_\omega$ is:

$$\begin{aligned} \frac{\partial \Delta e_\omega^*}{\partial x_s}(x) &= -t^* - \frac{\partial t^*}{\partial x_s} \left(\sum_{i \in \mathcal{B}_\omega} (\bar{g}_i - x_i) + \sum_{i \in \mathcal{B}_\omega^c} \rho_i t^* - \sum_{i \in \mathcal{O}_\omega} x_i \right) \\ &= -t^*. \end{aligned} \quad (\text{A.4})$$

Let $h(x) = \underline{E}_\omega - (E_\omega + \Delta e_\omega(x))$. Then (4.13) can be re-written as:

$$e_\omega^s(x) = \max(h(x), 0). \quad (\text{A.5})$$

In the case that $e_\omega^s(x) = h(x) > 0$, we have that

$$\frac{\partial e_\omega^s}{\partial x_s}(x) = \frac{\partial h}{\partial x_s}(x) = -\frac{\partial \Delta e_\omega^*}{\partial x_s}(x).$$

On the other hand, if $h(x) \leq 0$, $e_\omega^s(x) = 0$, which implies $\frac{\partial e_\omega^s}{\partial x_s}(x) = 0$. With the partial derivatives in (A.2)-(A.4) and definition of $\hat{p}_\omega^s(x)$ (4.16), we have $\beta_{i\omega}$, a subgradient of $\hat{p}_\omega^s(x)$, as follows:

$$\beta_{i\omega}(x) = \begin{cases} \gamma_\omega(t^* - t_s), & e_\omega^s(x) > 0, i \in \mathcal{B}_\omega(t^*), \\ 0, & e_\omega^s(x) = 0, i \in \mathcal{B}_\omega(t^*), \\ 0, & i \in \mathcal{B}_\omega^c(t^*), \\ \gamma_\omega t^*, & e_\omega^s(x) > 0, i \in \mathcal{O}_\omega \\ 0, & e_\omega^s(x) = 0, i \in \mathcal{O}_\omega. \end{cases} \quad (\text{A.6})$$

Note that $\Delta e_\omega^*(x)$ might be non-differentiable at some x , which occurs when some generators have just binding capacity constraints at t^* . In this case, any directional derivative can serve as a subgradient. One approach to evaluate a particular directional derivative is applying a small change in x_s and re-computing the corresponding t^* , $\mathcal{B}_\omega(t^*)$, and $\mathcal{B}_\omega^c(t^*)$. The directional derivative is then calculated using (A.2)-(A.4). Note that equations (A.2)-(A.4) also provide a subgradient of $\hat{p}_\omega^s(x)$ in the case of governor delay. The derivation is similar and therefore omitted.

Appendix B

Conditions for Preserving Optimality Cuts

For scenario ω with any first-stage decisions (x, r) , let γ'_ω and σ'_ω be initial values of γ_ω and σ_ω in (4.16), and let $Q'_\omega(x, r)$ be the associated second-stage problem. Similarly, define γ''_ω and σ''_ω as the updated values and let $Q''_\omega(x, r)$ be the associated second-stage problem. If:

$$\gamma''_\omega \geq \gamma'_\omega \text{ and } \sigma''_\omega \geq \sigma'_\omega, \tag{B.1}$$

then the right-hand side of (4.2b) of $Q''_\omega(x, r)$ is greater than or equal to that of $Q'_\omega(x, r)$. This implies $Q''_\omega(x, r)$ has a smaller feasible region and larger optimal value, that is,

$$Q''_\omega(x, r) \geq Q'_\omega(x, r).$$

Therefore, the existing optimality cuts, which are lower bounds of $Q'_\omega(\bullet)$, are still valid lower bounds of the updated second-stage problem $Q''_\omega(\bullet)$ and therefore can be preserved. The condition (B.1) also suggests that we can start with relatively smaller initial values, in order to avoid removing the optimality cuts after updating the coefficients.

Bibliography

- [1] B. Allaz and J.-L. Vila. Cournot competition, forward markets, and efficiency. *Journal of Economic Theory*, 59:1–17, 1993.
- [2] Argonne National Laboratory. A survey on wind power ramp forecasting, 2010. Available from <http://www.dis.anl.gov/pubs/69166.pdf>. Accessed 14 Feb 2012.
- [3] AWS Truewind. Analysis of west Texas wind plant ramp-up and ramp-down event, 2008. Available from http://interchange.puc.state.tx.us/WebApp/Interchange/Documents/33672_1014_580034.PDF. Accessed 14 Feb, 2012.
- [4] R. Baldick. Joint equilibrium of day-ahead and real-time markets. Unpublished manuscript, Department of Electrical and Computer Engineering, The University of Texas at Austin, Aug. 2002.
- [5] R. Baldick. Course notes for EE394V restructured electricity markets: locational marginal pricing, 2011. available from <http://users.ece.utexas.edu/~baldick/classes/394V/EE394V.html>. Accessed on 18 Apr. 2012.
- [6] J. Barquín. Symmetry properties of conjectural price responses. In *Proceedings of the IEEE Power Engineering Society General Meeting*, pages

1–7, Jul. 2008.

- [7] J. F. Benders. Partitioning procedures for solving mixed-variables programming problems. *Numerische Mathematik*, 4:238–252, 1962.
- [8] J. R. Birge and H. Tang. L-shape method for two stage problems of stochastic convex programming, 1993. Technical report, available from <http://deepblue.lib.umich.edu/bitstream/2027.42/3628/5/bbm0214.0001.001.pdf>. Accessed on 14 Feb 2012.
- [9] S. Borenstein, J. Bushnell, and C. R. Knittel. Market power in electricity markets: Beyond concentration measures. *The Energy Journal*, 20(4):68–88, 1999.
- [10] S. Borenstein, J. Bushnell, and S. Stoft. The competitive effects of transmission capacity in a deregulated electricity industry. *RAND Journal of Economics*, 31(2):294–325, Summer 2000.
- [11] F. Bouffard, F. Galiana, and A. Conejo. Market-clearing with stochastic security-part I: formulation. *IEEE Transactions on Power Systems*, 20(4):1818–1826, Nov. 2005.
- [12] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, Cambridge and New York, 2004.
- [13] J. Bushnell. Oligopoly equilibria in electricity contract markets. *Journal of Regulatory Economics*, 31:225–245, 2007.

- [14] California Independent System Operator. Competitive path assessment for MRTU, 2009. Available from <http://www.caiso.com/2365/23659ca314f0.pdf>. Accessed Jul. 2, 2010.
- [15] California Independent System Operator. Residual supply metrics: Methodology and preliminary 2009 results, 2010. Available from <http://www.caiso.com/2725/2725e3899550.pdf>. Accessed Jul. 2, 2010.
- [16] California Independent System Operator. Business practice manual: Market operations, 2011. Available from <https://bpm.caiso.com/bpm/bpm/version/000000000000157>. Accessed on 14 Feb 2012.
- [17] J. B. Cardell, C. C. Hitt, and W. W. Hogan. Market power and strategic interaction in electricity networks. *Resource and Energy Economics*, 19(1–2):109–137, Mar. 1997.
- [18] R. Doherty, G. Lalor, and M. O’Malley. Frequency control in competitive electricity market dispatch. *IEEE Transactions on Power Systems*, 20(3):1588–1596, Aug. 2005.
- [19] T. Dy Liacco. The adaptive reliability control system. *IEEE Transactions on Power Apparatus and Systems*, PAS-86(5):517–531, May 1967.
- [20] Eirgrid. Operating reserve requirements, 2010. Available from <http://www.eirgrid.com/aboutus/publications>. Accessed on 11 Feb 2012.
- [21] Electric Reliability Council of Texas. ERCOT protocols, 2011. Available from <http://www.ercot.com>. Accessed 22 Apr. 2012.

- [22] A. V. Fiacco and J. Kyparisis. Convexity and concavity properties of the optimal value function in parametric nonlinear programming. *Journal of Optimization Theory and Applications*, 48:95–126, 1986.
- [23] GE Energy. Analysis of wind generation impact on ERCOT ancillary services requirements, 2008. Available from http://www.uwig.org/AttchA-ERCOT_A-S_Study_Exec_Sum.pdf. Accessed 14 Feb 2012.
- [24] T. Genc and S. S. Reynolds. Supply function equilibria with pivotal suppliers. Unpublished manuscript, March 2005.
- [25] A. M. Geoffrion. Generalized Benders decomposition. *Journal of Optimization Theory and Applications*, 10(4):237–260, 1972.
- [26] B. F. Hobbs, C. Metzler, and J.-S. Pang. Strategic gaming analysis for electric power networks: An MPEC approach. *IEEE Transactions on Power Systems*, 15(2):638–645, May 2000.
- [27] W. W. Hogan. A market power model with strategic interaction in electricity markets. *The Energy Journal*, 18(4):107–141, 1997.
- [28] W. W. Hogan. Connecting reliability standards and electricity markets, 2005. Available from <http://www.hks.harvard.edu/fs/whogan>. Accessed on 14 Feb 2012.
- [29] W. W. Hogan. A model for zonal operating reserve demand curve, 2009. Available from <http://www.hks.harvard.edu/fs/whogan>. Accessed on 12 Feb 2012.

- [30] T. Inoue, H. Taniguchi, Y. Ikeguchi, and K. Yoshida. Monitoring the first frequency derivative to improve adaptive underfrequency Load-Shedding schemes. *IEEE Transactions on Power Systems*, 12(1):136–143, May 1997.
- [31] ISO New England. ISO New England operating procedure No. 8, 2011. Available from http://www.iso-ne.com/rules_proceeds/operating/isone/op8/op8_rto_final.pdf. Accessed on 14 Feb 2012.
- [32] J. J. E. Kelley. The cutting-plane method for solving convex programs. *Journal of the Society for Industrial and Applied Mathematics*, 8(4):703–712, Dec. 1960.
- [33] P. L. Joskow. California’s electricity crisis. MIT, Sep. 2001.
- [34] E. Kahn and R. Baldick. Reactive power is a cheap constraint. *The Energy Journal*, 15(4):191–201, 1994.
- [35] L. Kimball, K. Clements, and P. Davis. Stochastic OPF via Bender’s method. In *Power Tech Proceedings, 2001 IEEE Porto*, volume 3, page 4 pp. vol.3, 2001.
- [36] Lawrence Berkeley National Laboratory. Use of frequency response metrics to assess the planning and operating requirements for reliable integration of variable renewable generation, 2010. Available from <http://www.ferc.gov>. Accessed on 14 Feb 2012.

- [37] B. C. Lesieutre, K. M. Rogers, T. J. Overbye, and A. R. Borden. A sensitivity approach to detection of local market power potential. *IEEE Transactions on Power Systems*, 26(4):1980–1988, Nov. 2011.
- [38] B. C. Lesieutre, R. J. Thomas, and T. D. Mount. Identification of load pockets and market power in electric power systems. *Decision Support Systems*, 40:517–528, Oct. 2005.
- [39] MathWorks Corporation. Matlab, 2012. Available from <http://www.mathworks.com/>. Accessed 18 Apr. 2012.
- [40] Midwest Independent System Operator. MISO business practices manual: Energy and operating reserve markets, 2010. Available from <http://www.midwestiso.org>. Accessed on 14 Feb 2012.
- [41] C. Murillo-Sanchez, S. Ede, T. Mount, R. Thomas, and R. Zimmerman. An engineering approach to monitoring market power in restructured markets for electricity. In *Proceedings of the 24th Annual International Conference, International Association for Energy Economics*, Apr. 2001.
- [42] New York Independent System Operator. Ancillary services manual, 2010. Available from <http://www.nyiso.com/public/webdocs/documents/manuals/operations/ancserv.pdf>. Accessed Sep. 21, 2010.
- [43] D. M. Newbery. Competition, contracts, and entry in the electricity spot market. *RAND Journal of Economics*, 29(4):726–749, Winter 1998.

- [44] D. M. Newbery. Predicting market power in wholesale electricity markets. Cambridge University EPRG working paper 0821, Sep. 2008.
- [45] North America Electric Reliability Corporation. Background: industry advisory. Reliability risk interconnection frequency response. (Revision 1), 2010. Available from <http://www.nerc.com/page.php?cid=5163>. Accessed on 24 Apr. 2012.
- [46] North America Electric Reliability Corporation. Generating availability data system (GADS), 2010. Available from <http://www.nerc.com>. Accessed on 14 Feb 2012.
- [47] North America Electric Reliability Corporation. Real-time application of synchrophasors for improving reliability, 2010. Available from <http://www.nerc.com/docs/oc/rapirtf/RAPIR%20final%20101710.pdf>. Accessed on 24 Apr. 2012.
- [48] North American Electric Reliability Corporation. Standard BAL-002-0-disturbance control performance, 2005. Available from <http://www.nerc.com/files/BAL-002-0.pdf>. Accessed on 14 Feb 2012.
- [49] A. Papavasiliou, S. S. Oren, and R. P. O'Neill. Reserve requirements for wind power integration: A scenario-based stochastic programming framework. *IEEE Transactions on Power Systems*, 26(4):2197–2206, Nov. 2011.

- [50] PJM Interconnection. 2006 state of the market report, volume ii: Detailed analysis, March 2007. Available from http://www.monitoringanalytics.com/reports/PJM_State_of_the_Market/2006.shtml. Accessed on 10 Apr 2012.
- [51] PJM Members Committee. Reliability pricing model overview. Available from <http://www.pjm.com>, 2004.
- [52] PowerWorld Corporation. PowerWorld contouring and advanced visualization, 2008. available from <http://www.powerworld.com/WebTraining/I06ContouringAdvancedVisualization.pdf>. Accessed 18 Apr. 2012.
- [53] PowerWorld Corporation. PowerWorld Simulator, 2012. available from <http://www.powerworld.com>. Accessed 18 Apr. 2012.
- [54] Reliability Test System Task Force. The IEEE reliability test system—1996. *IEEE Transactions on Power Systems*, 14(3):1010–1020, Aug. 1999.
- [55] J. F. Restrepo and F. D. Galiana. Unit commitment with primary frequency regulation constraints. *IEEE Transactions on Power Systems*, 20(4):1836–1842, Nov. 2005.
- [56] P. Ruiz, C. Philbrick, E. Zak, K. Cheung, and P. Sauer. Uncertainty management in the unit commitment problem. *IEEE Transactions on Power Systems*, 24(2):642–651, May 2009.

- [57] P. Ruiz and P. Sauer. Spinning contingency reserve: Economic value and demand functions. *IEEE Transactions on Power Systems*, 23(3):1071–1078, Aug. 2008.
- [58] A. Saric, F. Murphy, A. Soyster, and A. Stankovic. Two-stage stochastic programming model for market clearing with contingencies. *IEEE Transactions on Power Systems*, 24(3):1266–1278, Aug. 2009.
- [59] A. Y. Sheffrin and J. Chen. Predicting market power in wholesale electricity markets. In *Proceedings, 15th Annual Western Conference of the Advanced Workshop in Regulation and Competition, South Lake Tahoe, CA*, 2002.
- [60] A. Y. Sheffrin, J. Chen, and B. F. Hobbs. Watching watts to prevent abuse of power. *IEEE Power and Energy Magazine*, 2(4):58–65, Jul./Aug. 2004.
- [61] R. M. V. Slyke and R. Wets. L-shaped linear programs with applications to optimal control and stochastic programming. *SIAM Journal on Applied Mathematics*, 17(4):638–663, 1969.
- [62] C. B. Somuah and F. C. Schweppe. Economic dispatch reserve allocation. *IEEE Transactions on Power Apparatus and Systems*, PAS-100(5):2635–2642, May 1981.
- [63] Southwest Power Pool Electric Energy Network. SPP market protocols, 2010. Available from <http://www.spp.org>. Accessed on 14 Feb 2011.

- [64] S. Stoft. *Power System Economics: Designing Markets for Electricity*. IEEE Press and Wiley Interscience and John Wiley & Sons, Inc., Piscataway, NJ, 2002.
- [65] S. Takriti, J. Birge, and E. Long. A stochastic model for the unit commitment problem. *IEEE Transactions on Power Systems*, 11(3):1497–1508, Aug. 1996.
- [66] Transpower. Reserve management tool, 2004. Available from <http://www.electricitycommission.govt.nz/opdev/servprovinform/servprovpdfs/spc-system-operator-rmt-specification.pdf>. Accessed Sep. 21, 2010.
- [67] A. Tuohy, E. Denny, P. Meibom, R. Barth, and M. O’Malley. Operating the Irish power system with increased levels of wind power. In *IEEE Power and Energy Society General Meeting - Conversion and Delivery of Electrical Energy in the 21st Century*, pages 1–4, Jul. 2008.
- [68] H. R. Varian. *Intermediate Microeconomics: a Modern Approach*. W. W. Norton & Company, New York, seventh edition, 2005.
- [69] F. A. Wolak. Identification and estimation of cost functions using observed bid data, 2002. Available from <http://www.stanford.edu/wolak>. Accessed on 14 Oct. 2010.
- [70] F. A. Wolak. Measuring unilateral market power in wholesale electricity markets: The California market 1998–2000. *The American Economic Review*, 93(2):425–430, May 2003.

- [71] A. J. Wood and B. F. Wollenberg. *Power Generation, Operation, and Control*. Wiley, New York, second edition, 1996.
- [72] L. Xu and R. Baldick. Transmission-constrained residual demand derivative in electricity markets. *IEEE Transactions on Power Systems*, 22(4):1563–1573, Nov. 2007.
- [73] L. Xu, R. Baldick, and Y. Sutjandra. Transmission-constrained inverse residual demand Jacobian matrix in electricity markets. *IEEE Transactions on Power Systems*, 26(4):2311–2318, Nov. 2011.
- [74] C. Yonghong and P. Gribik. Incorporating reserve deployment impact on transmission constraints into co-optimization of energy and ancillary service procurement. In *FERC Technical Conference on Increasing Real-Time and DayAhead Market Efficiency through Improved Software*, Jun. 2011.

Vita

Yen-Yu Lee was born in Kaohsiung, Taiwan on 3 May 1984. He attended National Taiwan University and received the B.S. degree in Electrical Engineering in 2006. He worked as a research assistant at Academia Sinica from 2006 to 2007. He then attended the University of Texas at Austin and received the M.S. in electrical and computer engineering in 2010. He worked as a summer intern at Cadence Design System, Pyxis Technology, and the Electric Reliability Council of Texas.

Permanent address: 13F No.8 Sin-Guo Rd Kaohsiung Taiwan 807

This dissertation was typeset with L^AT_EX[†] by the author.

[†]L^AT_EX is a document preparation system developed by Leslie Lamport as a special version of Donald Knuth's T_EX Program.