Decomposition of Variational Inequalities with Applications to Nash-Cournot Models in Time of Use Electricity Markets

by

Emre Çelebi

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Management Sciences

Waterloo, Ontario, Canada, 2011

AUTHOR'S DECLARATION

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

ABSTRACT

This thesis proposes equilibrium models to link the wholesale and retail electricity markets which allow for reconciliation of the differing time scales of responses of producers (e.g., hourly) and consumers (e.g., monthly) to changing prices. Electricity market equilibrium models with time of use (TOU) pricing scheme are formulated as large-scale variational inequality (VI) problems, a unified and concise approach for modeling the equilibrium. The demand response is dynamic in these models through a dependence on the lagged demand. Different market structures are examined within this context. With an illustrative example, the welfare gains/losses are analyzed after an implementation of TOU pricing scheme over the single pricing scheme. An approximation of the welfare change for this analysis is also presented. Moreover, break-up of a large supplier into smaller parts is investigated.

For the illustrative examples presented in the dissertation, overall welfare gains for consumers and lower prices closer to the levels of perfect competition can be realized when the retail pricing scheme is changed from single pricing to TOU pricing. These models can be useful policy tools for regulatory bodies i) to forecast future retail prices (TOU or single prices), ii) to examine the market power exerted by suppliers and iii) to measure welfare gains/losses with different retail pricing schemes (e.g., single versus TOU pricing).

With the inclusion of linearized DC network constraints into these models, the problem size grows considerably. Dantzig-Wolfe (DW) decomposition algorithm for VI problems is used to alleviate the computational burden and it also facilitates model management and maintenance. Modification of the DW decomposition algorithm and approximation of the DW master problem significantly improve the computational effort required to find the equilibrium. These algorithms are applied to a two-region energy model for Canada and a realistic Ontario electricity test system. In addition to empirical analysis, theoretical results for the convergence properties of the master problem approximation are presented for DW decomposition of VI problems.

ACKNOWLEDGEMENTS

It is not overstating that this thesis would not have been possible without my supervisor, Dr. J. David Fuller. Dave has been a great mentor to me with his guidance, encouragement, constructive critiques, seemingly boundless patience, insightful interest, modesty and enthusiasm throughout my graduate studies. I have not only completed this thesis with him, but learned a lot from him about how to be an excellent and successful academic.

I would like to thank to my committee members, Dr. Samir Elhedhli, Dr. Miguel Anjos, Dr. Kankar Bhattacharya and Dr. Richard O'Neill for their valuable comments, insights and feedback. I am also thankful to Dr. William Chung and Dr. Steven Wong for providing me the data and models for the illustrations in the thesis. I appreciate the generous funding from Natural Sciences and Engineering Research Council of Canada (NSERC) and Dept. of Management Sciences, University of Waterloo (UW).

After eight years of graduate studies at UW, I am indebted to many good friends and colleagues that have inspired, encouraged and supported me. I am especially thankful to Serdar Tarı, Mary-Helen & M. Ali Ülkü, Özgür Çelikoğlu, Mehmet Can Tarı, Canan Akdemir, Mehmet Gümüş, Selvaprabu Nadarajah, Christine & Selçuk Onay, Aşan Bacak, Navneet Vidyarthi, Laura & Andrei Radulescu, Joe Naoum-Sawaya, Bissan Ghaddar, Güneş Aluç, Burcu & Koray Karabina, Bengücan & Onur Canpolat, Doruk Babaoğlu, Hakan Öztürkoğlu, Emre Karakoç, Gordon Adomdza, Mark Zschocke, Sachin Jawasal, Ada & Adam Hurst, Mehrdad Pirnia, Haldun Süral, James Bookbinder, Frank Safayeni and Walaa Elshabrawy for their devoted and lasting companionship as well as for enjoyable times together.

I am grateful to my sisters and their families, Elif & Cevdet & Berfin Naz, Pınar & Yücel & Ege Eren, who have never let me feel alone. My parents, Necla and Habib Çelebi, have encouraged, supported, understood, and loved me at every moment. None of this could have happened without them. This thesis is dedicated to them.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ACCPM: Analytic Center Cutting Plane Method

CG: Convergence Gap

CPU: Central Processing Unit

DC: Direct Current

DW: Dantzig-Wolfe

EMP: Extended Mathematical Programming

GAMS: General Algebraic Modeling System

IESO: Independent Electricity System Operator

ISO: Independent System Operator

KKT: Karush-Kuhn-Tucker

LP: Linear Programming

MCP: Mixed Complementarity Problem

NEMS: National Energy Modeling System

NLP: Nonlinear Programming

OEB: Ontario Energy Board

OPG: Ontario Power Generation

PCM: Price Cost Margin

PJM: Pennsylvania New Jersey Maryland

PIES: Project Independence Evaluation System

PTDF: Power Transfer Distribution Factor

RCG: Relative Convergence Gap

RTO: Regional Transmission Operator

RTP: Real Time Pricing

TOU: Time of Use

VI: Variational Inequality

LIST OF SYMBOLS

This is a non-exhaustive list of commonly used symbols throughout the thesis. Subscript n represents the nodes (e.g., buses) and the superscript *(t)* denotes the periods (e.g., month).

Sets

set of periods (months): $t=1,\ldots,T$
set of generation facilities: $i=1,\ldots,I$
set of demand blocks: $j=1,\ldots,J$ (alias index k)
set of nodes: $n=1,\ldots,N$ (N_d: set of demand nodes; N_g: set of generation nodes)
set of hours in a demand block j: $h=1,\ldots,H_j$ (defined by the market regulator)
set of firms: $f=1,\ldots,F$
set of lines: $l=1,..,L$

Parameters

c_{fi} = operating cost per unit of energy for firm f's facility i (\$/MWh)
κ_{fi} = capacity of firm f's facility i (MW)
δ_{jh} = fraction of total energy demand during block j of a month for hour h
$PTDF_{ln}$ = power transfer distribution factors
T_{l-}, T_{l+} = lower and upper bounds on real power flows through line (interface) l (MW)
$\boldsymbol{A}$ = vector of the factors representing non-price effects
$\boldsymbol{B}$ = a square matrix of the price coefficients (i.e., own-price and cross-price)
$\boldsymbol{E}$ = a square diagonal matrix of the lag coefficients
$\boldsymbol{D}^0$ = vector of lagged demands

Decision variables

z_{fijh} = the energy flowing from firm f's facility i to demand block j for hour h (MWh)
d_{fj} = sales by firm f to demand block j (MWh)
p_j = TOU prices (e.g., uniform block prices) for demand block j (off-peak, mid-peak and on-peak, \$/MWh)
y_{njh} = net injections from transmission lines into node n for demand block j at hour h (MW)
β_{njh} = a congestion based fee (e.g., wheeling fee) for transmitting power from an arbitrary hub node to node n.

1. Introduction

The desire to solve more and more complex models has always been running ahead of capabilities of the time, and has provided a driving force for the development of faster and reliable algorithms to compute a solution. Complex models arise in many areas such as engineering, economics, public policy analysis and energy markets. They are defined as "one made of a large number of parts that interact in a non-simple way" and complexity of a model increases by the number of different model parts, number of different ways these parts can interact, number of different type of variables and functions, time and spatial dimension and by the scope and range of the problem environment (Murphy, 1993).

Complex and large-scale equilibrium models have been at the heart of research interest for many researchers and practitioners. Advances in computation power and increase in availability of data in the last few decades also has supported a dramatic growth of interest in these models. They have removed the need for working on small dimensional analytical models and incorporate much more detail than analytical models. This has stimulated the use of different techniques or algorithms to solve such equilibrium models, e.g., the Project Independence Evaluation System (PIES) algorithm (Ahn and Hogan, 1982), the decoupling algorithm (Wu and Fuller, 1995), and algorithms for complementarity problems (Mathiesen, 1985; Manne, 1985; Dirkse and Ferris, 1996; Ferris et al., 2001) or VI problems (Nagurney, 1993). The former two algorithms are based on a sequence of integrable optimization problems, whereas the latter two are more general and recognized approaches.

PIES, which was originally developed for energy modeling for U.S. Department of Energy in the 1970s, captures many key features of complex equilibrium models. It has proven its success by solving such equilibrium models. The PIES algorithm approximates the non-integrable equilibrium problem by a sequence of integrable problems which can be converted into equivalent optimization problems. Each iteration solves a linear programming (LP) problem after a proper step function approximation is made on an integrable approximation of the demand function (Hogan, 1975). This algorithm has the characteristics of the nonlinear Jacobi method for solving nonlinear system of equations. Ahn and Hogan (1982) give sufficient conditions under which the PIES algorithm converges. But, as Murphy and Mudrageda (1998) point out, although PIES never met these conditions, because of demand function approximations, it usually does not fail to converge.

Variational inequality (VI) problems were first developed in the context of studying a class of partial differential equations that arise in the field of mechanics and defined on infinite dimensional spaces. In contrast, finite dimensional VI problems have been studied for computation of economic and game theoretic equilibria. For a comprehensive survey see Harker and Pang (1990). Many mathematical problems (e.g., system of equations, constrained and unconstrained optimization problems, complementarity problems, game theory and saddle point problems, fixed point problems, traffic assignment and network equilibrium problems) can be formulated as VI problems (Nagurney, 1993; Harker and Pang, 1990; Bertsekas and Tsitsiklis, 1989; Patriksson, 1994).

There are several decomposition algorithms (e.g., Dantzig-Wolfe, Benders, Lagrangean) for solving and analyzing complex models. Decomposition is a notion to split a mathematical program into two or more set of variables and associated constraints. Separating some parts with special structure from the rest of the mathematical programming model is the purpose of the decomposition. Certain models may have a structure that some of the constraints prevent the separability of the problem into subproblems. If these constraints are removed, the resulting subproblems are frequently considerably easier to solve. These constraints are usually referred to as "complicating constraints" (and sometimes referred to as "common" or "linking" constraints). Dantzig-Wolfe (DW) decomposition is developed to take advantage of this structure (e.g., block diagonal structure). Some other models may have a structure that if certain variables (i.e., binary and/or integer) can be fixed, the problem reduces to one that is easier to solve. These models including such variables are referred to as problems with “complicating variables”. Benders decomposition is developed to exploit the structure of such models (Conejo et al., 2006).

Decomposition methods allow large-scale and complex problems to be solved in a distributed and parallel fashion that helps to overcome computational difficulties (Fuller and Chung, 2008). They can reduce the memory requirements and/or increase the speed of calculations. Alternatively they can lead to a drastic simplification of the model development procedure and ease the model management and maintenance (Murphy, 1993; Murphy and Mudrageda, 1998). Generally, the scope of the complex models (e.g., related to public policy making) expands as addressing one question reveals other related questions. Therefore analyses of such models require continuous re-evaluation of the issues. Decomposition of these models allows different analysts or teams of experts to manage, analyze, re-evaluate and repeatedly run sub-models. Murphy (1993) demonstrated that expected run time can be reduced by 70% using decomposition methods. Another advantage of decomposition methods pointed out by Murphy (1993) is the error reduction in modeling (e.g., effects of errors can be made additive rather than multiplicative).

Any decomposition algorithm involves an iterative procedure to solve subproblems by using adjusted information passed to subproblems between iterations. One type of algorithm is to use a “master problem” to coordinate the communication between subproblems and adjust the information for the subproblems (Fuller and Chung, 2008). Instead of solving the original problem with complicating constraints or variables, two problems are solved iteratively, a master problem and a subproblem, i.e., original problem without complicating constraints or variables (it can also be decomposed into several subproblems in the absence of complicating constraints or variables). The solution to the original model is obtained by exchanging price and quantity information among the subproblem(s) and the master problem in an iterative manner.

The power of formulation with VI framework extends our interest in the areas of very well studied algorithms and techniques for optimization problems, mainly decomposition, to alleviate the difficulties when confronted with large-scale and complex models, (e.g., multiregional economic equilibrium problems, multiplayer game theoretic electricity market models). Many equilibrium problems can be formulated as VI problems. But it is a challenging task to find and to analyze the equilibrium for large-scale and complex models. However, such models can have a special structure, i.e., they can be decomposable by some dimension (e.g., time, region, sector, product, player). Therefore, we are motivated to apply decomposition algorithms from optimization theory to VI problems and analyze the performance of such algorithms with further modifications and testing on large-scale, realistic models.

The growing literature on market equilibrium models of electric power networks under different market designs presents examples of large-scale and complex equilibrium problems (Helman and Hobbs, 2010). The network underlying such market equilibrium problems provides an added dimension to the analysis and computation of equilibria (due to network constraints governed by Kirchhoff's current and voltage laws). Simulation of market pricing through equilibrium market models can inform policy makers and system operators (e.g., regulatory bodies and independent system/market operators) by providing insights about the market design and market power issues (e.g., merger/divestiture decisions) (Smeers, 1997; Wei and Smeers, 1999; Day et al., 2002). Moreover, fairly detailed models can be useful in short-/long-term forecasting and ex-post replication of prices (Green and Newberry, 1992; Borenstein and Bushnell, 1998).

Many models consider electricity as an undifferentiated commodity (e.g., a single product). However, in actual electricity markets, electricity is differentiated by both time (e.g., time of use –TOU) and location (e.g., network node) and even the fuel type (e.g., fossil fuel/nuclear versus renewable energy) (Hobbs and Pang, 2004). In this thesis, we propose equilibrium models of competition with product differentiation (e.g., using TOU pricing) as a more adequate approach to analyze strategic behavior in electricity markets. For a better understanding of the market behavior and outcomes, such models can be used in regulatory procedures (e.g., a guide for market power monitoring and mitigation, merger/divestiture analysis) and evaluation/forecasting of prices (e.g., forecasting forward prices in markets where there is no consumer price regulation).

The proposed models in this thesis also address the problem of demand response with the inclusion of TOU pricing scheme in the context of bilateral markets. Our focus in this thesis remains in, but not limited to, the Ontario electricity market. Ontario has been one of the pioneers in TOU pricing scheme and invested in smart metering technologies substantially. The aim is to provide an incentive for consumers to shift some of their consumption away from periods of high demand to periods of low demand and conserve often by installing more energy efficient equipment. This move in Ontario towards TOU pricing will certainly affect the market structure in the future. Some market participants may be interested in TOU bilateral contracts and in the future they can form a TOU bilateral market for Ontario. In such a bilateral market, power quantities for predefined intervals of hours (off-peak, mid-peak and on-peak) at negotiated prices, terms and conditions can be traded among participants. Within this context, the proposed models can be used as forecasting and policy analysis tools.

1.1 Variational Inequality Problems

The models proposed in this thesis are represented and solved by the VI problem approach. In general, a finite dimensional VI problem is defined as follows:

$$\begin{gathered} VI(\boldsymbol{F},K)\text{: Find a vector } \boldsymbol{x}^* \in K \subset R^n\text{, such that:} \\ \boldsymbol{F}(\boldsymbol{x}^*)^T.(\boldsymbol{x}-\boldsymbol{x}^*) \geq 0, \quad \forall \boldsymbol{x} \in K \end{gathered} \tag{1.1}$$

where $\boldsymbol{F}$ is a given continuous function from K to R^n and K is a nonempty, closed and convex set (Nagurney, 1993). Unlike an optimization problem which has an objective function, a VI problem has a vector-valued function $\boldsymbol{F}$, and it is equivalent to an optimization problem only if this vector-valued function is the gradient of an objective function. A necessary and sufficient condition for a differentiable $\boldsymbol{F}$ to satisfy the above condition is that the Jacobian matrix $\nabla\boldsymbol{F}$ is symmetric or in other words, that $\boldsymbol{F}$ is integrable, i.e., it can be integrated to define an objective function (Nagurney, 1993).

Unfortunately, this condition does not hold in many practical problems. In this thesis, we consider problems which are non-integrable (asymmetric)[1].

Standard conditions for existence and uniqueness of solutions to $VI(\boldsymbol{F}, K)$ are provided below. Definitions for monotonicity and coercivity are also provided (Patriksson, 1994). See Harker and Pang (1990), Nagurney (1993) or Patriksson (1994) for proofs.

Definition 1.1 (Monotonicity):

1) $\boldsymbol{F}$ is *monotone* on K if $[\boldsymbol{F}(\boldsymbol{x}) - \boldsymbol{F}(\boldsymbol{y})]^T.(\boldsymbol{x} - \boldsymbol{y}) \geq 0, \quad \forall \boldsymbol{x}, \boldsymbol{y} \in K$.

2) $\boldsymbol{F}$ is *strictly monotone* on K if $[\boldsymbol{F}(\boldsymbol{x}) - \boldsymbol{F}(\boldsymbol{y})]^T.(\boldsymbol{x} - \boldsymbol{y}) > 0, \quad \forall \boldsymbol{x}, \boldsymbol{y} \in K, \ \boldsymbol{x} \neq \boldsymbol{y}$.

3) $\boldsymbol{F}$ is *strongly (uniformly) monotone* on K if there exists a positive constant $m_{\boldsymbol{F}}$ such that $[\boldsymbol{F}(\boldsymbol{x}) - \boldsymbol{F}(\boldsymbol{y})]^T.(\boldsymbol{x} - \boldsymbol{y}) \geq m_{\boldsymbol{F}} \|\boldsymbol{x} - \boldsymbol{y}\|^2, \quad \forall \boldsymbol{x}, \boldsymbol{y} \in K$.

Definition 1.2 (Coercivity): $\boldsymbol{F}$ is *coercive* on K if there exists a vector $\boldsymbol{x}^0 \in K$ such that

$$\lim_{\substack{\boldsymbol{x} \in K \\ \|\boldsymbol{x}\| \to +\infty}} \frac{\left(\boldsymbol{F}(\boldsymbol{x}) - \boldsymbol{F}(\boldsymbol{x}^0)\right)^T.(\boldsymbol{x} - \boldsymbol{x}^0)}{\|\boldsymbol{x} - \boldsymbol{x}^0\|} = +\infty.$$

Theorem 1.1: The existence of the solution to a VI problem follows from any one of the following properties of $\boldsymbol{F}$ or K:

1) The set K is bounded;

2) The mapping $\boldsymbol{F}$ is strongly monotone;

---

[1] Samuelson (1951) formulated a spatial equilibrium model as a mathematical programming problem with the objective of maximizing consumers' plus producers' surpluses (e.g., "net social payoff") in different regional markets minus the transportation costs. Takayama and Judge (1971) extended Samuelson's model using a "quasi welfare function" with linear price dependent demand and supply functions. They also defined the "integrability" conditions to form this quasi welfare function and proposed a quadratic programming algorithm to obtain the equilibrium solution. But, if the demand functions are not symmetric, the integrability conditions are not satisfied. For example, in a multi-commodity model, cross price effects in demand can cause the integrability conditions to fail. Empirically estimated demand functions are unlikely to satisfy these conditions. However, VI problems overcome this shortcoming of optimization approaches.

3) The mapping $\boldsymbol{F}$ is coercive.

Theorem 1.2: The solution of a VI problem, if one exists, is unique, if $\boldsymbol{F}$ is strictly monotone on K.

Complementarity problems are considered as a special case of VI problems, where they are defined on the non-negative orthant (proof can be found in Nagurney, (1993), p.7-8). Many electricity market models are represented using the complementarity problem framework. For examples see Hobbs and Helman (2010) and references therein with the exception of Wei and Smeers (1999), Daxhelet and Smeers (2001) who use VI framework.

When Nash equilibrium (or generalized Nash equilibrium) problems are formulated as VI problems, they only include primal variables and their feasibility conditions, whereas complementarity problems include both primal and dual variables, their feasibility conditions together with complementary slackness conditions (i.e., more variables and conditions need to be represented and coding effort in GAMS is laborious compared to a VI problem formulation). It is possible to represent several objectives of different agents with VI formulation (e.g., in Nash equilibrium models), as we shall illustrate in chapters 2 and 3.

Recently, the extended mathematical programming (EMP) framework (available under GAMS) has made it possible to manage and maintain VI problems. The modeler is not required to formulate an equilibrium problem as a mixed complementarity problem (MCP) or a system of nonlinear equations (i.e., no need to represent dual variables and their feasibility conditions) (Ferris et al., 2009). The EMP framework automates the conversion of a VI problem into an equivalent MCP, which is then solved by the MCP solver PATH.

1.2 *Review of Time of Use (TOU) Pricing in Electricity Markets*

In this subsection we briefly give an overview of TOU literature. This subsection is mainly extracted from Celebi (2005). In economic theory, efficient pricing is achieved when electricity is priced at the marginal cost of supplying the last increment of electricity demand, and a perfectly competitive market can provide this. The literature on real time "Marginal Cost Pricing" in electricity markets is now vast (see the seminal work by Schweppe et al. (1988)). However, the time-differentiated pricing concept started earlier with studies on "Peak Load Pricing" (see Celebi and Fuller (2007) and references therein). Peak loads and their pricing have been a concern because of the capacity requirements for these loads. In peak load pricing, high marginal cost of electricity during the periods of the peak load is reflected in consumer prices, e.g., by time of use (TOU) pricing. In TOU pricing, both prices and time periods are known ex ante and are fixed for some duration (e.g., a season). In contrast, in real time pricing (RTP), generally prices change on an hourly basis and are fixed and known only on a day-ahead or hour-ahead basis. RTP reflects the wholesale prices, weather conditions, generator failures, scarcity of generation or other contingencies that occur in a wholesale electricity market.

The theoretical body of peak load pricing literature was not able to give practical answers to the problem and many large-scale experiments have been conducted with TOU pricing over the past three decades. Surveys of these experiments can be found in Mitchell et al. (1978), Faruqui and Malko (1983), Aigner (1984), King and Chatterjee (2003). These experiments collected data that allow econometricians to estimate the parameters of electricity demand functions such as own and cross price elasticities, elasticities of substitution and lag elasticities. Some countries even implemented TOU pricing on a national scale (see Mitchell et al. (1978) and Chick (2002) for an account). A more recent experiment for California (Statewide Pricing Pilot) has shown that residential and small to medium commercial and industrial customers cut energy usage in peak periods in response to TOU prices (Faruqui and George, 2005).

Patrick and Wolak (2001) estimated some significant demand response for electricity by large and medium-sized industrial and commercial customers purchasing electricity according to the half-hourly real time price from the England & Wales electricity market. Borenstein et al. (2002) reported on some U.S. utilities' implementations of RTP for their medium/large industrial and commercial customers that showed reduction in peak period energy use.

The main attraction of time varying pricing (e.g., RTP, TOU) is that it motivates customers to reduce their electricity consumption in peak periods and shifts some to off-peak periods, thereby reducing the total capacity requirement (Celebi and Fuller, 2007). However, RTP transfers the uncertainty and volatility of prices to customers and consequently, this has failed to attract many customers (Faruqui and George, 2002). Joskow and Tirole (2004) gave two main reasons why the final consumers may not react to real time prices. Firstly, the cost of monitoring and evaluation of hourly prices and constantly optimizing the use of equipment is enormous for small consumers. Secondly, most directed interruptions (due to a shortage in supply) that can be controlled by the distribution network operator usually occurs at the level of zones, which means that individual small consumers cannot sign up for interruptible supply contracts. For these reasons, TOU pricing is considered to be a practical step towards time-differentiated pricing.

1.3 Summary of Contributions

In this thesis, our main contributions can be summarized as follows:

a) We formulate market equilibrium models of competition with product differentiation (e.g., using TOU pricing) as a more adequate approach to analyze strategic behavior in electricity markets. These models allow for different pricing schemes (e.g., TOU or single or RTP) to be determined in different market structures. We develop policy tools for regulatory bodies to assess market power, to forecast (seasonal or monthly) TOU or single prices for future periods/seasons.
b) We approximate consumers' surplus for the consumers' habit formation model (i.e., using distributed lagged demand model) and present an approximate measure of welfare among other measures for regulatory bodies (e.g., market concentration index and price-cost margins –PCM). The welfare measure can be useful for analyzing the regulatory applications (e.g., regulated retail rates).

c) The models in the thesis are formulated by VI problem framework, which provides compact and convenient representation of equilibrium problems. Moreover, VI models are easier to manage, maintain and solve in GAMS/EMP framework than GAMS/MCP models. We have found some computational advantages for large-scale VI formulation over MCP formulations (e.g., improvements in computation times).
d) A modified DW algorithm is offered for large-scale market equilibrium models of TOU pricing with linearized network constraints. Exact DW algorithm is modified in a way that the independent system operator's (ISO's) problem is contained in the master problem and extra constraints are added to the subproblem. The modifications produce better proposals from subproblem to be passed to DW master problem. Subproblem without network constraints is easier to solve, build and manage, while the network constraints are contained in master problem. Without any line limits, this modification improves the DW algorithm substantially over the exact DW algorithm. Also, it allows different teams or analysts for master problem and subproblem to maintain and manage the model.
e) With line limits case, we propose an approximation of the master problem in the modified DW algorithm to gain computational advantages. We generalize this method for the DW decomposition with convergence analysis. Numerical results also support that the approximation of the master problem can overcome problems when confronted with computational limits (e.g., time and memory limits).

1.4 Overview of the Thesis

The thesis is organized as follows. Chapter 2 introduces the TOU pricing models in electricity markets under different market structures and their underlying assumptions. An illustrative example is presented, including several analyses to compare different market structures and pricing schemes (single versus TOU pricing, oligopoly versus perfect competition, break-up of a large firm into two or more parts with comparative welfare analyses). In chapter 3, linearized DC network is introduced into the TOU pricing model and VI formulations are provided. Chapter 4 provides an overview of the decomposition methods for VI problems. It presents applications of the DW decomposition algorithm to TOU pricing models and provides convergence analysis. Moreover, modifications to the algorithm and approximation of the master problem solution to gain computational advantages are provided with convergence analysis. The paper concludes with chapter 5, in which the directions for future research are suggested.

2. Time of Use (TOU) Pricing Models in Electricity Markets under Different Market Structures and Welfare Analysis[2]

2.1 *Introduction*

In the deregulation of many electricity markets, only wholesale markets have been open to competition and often, regulatory bodies have immunized retail markets against the price volatility and spikes in the wholesale markets by regulating consumer prices (e.g., Ontario regulated price plan). Except for large industrial customers equipped with real time meters, participation of the demand side in the wholesale market is very small. Consequently, many consumers are indifferent to fluctuations in electricity prices or uninterested in curtailing power usage during price spikes in the wholesale markets. Moreover, this does not encourage consumers to reduce peak demand, thereby causing supply costs to increase due to extra peak generation capacity[3].

The regulation of retail rates limits the price-response by many consumers to wholesale markets (Bompard et al., 2007), but there are some jurisdictions that seek to revive price responsiveness of consumers by offering time-differentiated regulated pricing schemes (e.g., TOU pricing in Ontario) rather than single pricing schemes. With TOU pricing, electricity prices are differentiated by time (e.g., hour of the day) and this can lead to models of competition with product differentiation as a more adequate

[2] This chapter is a journal submission under fourth round review: Celebi, E. and Fuller, J. D. 2011. "Time of Use Pricing in Electricity Markets under Different Market Structures," IEEE Transactions on Power Systems.

[3] Demand response is not only about price spikes but also about increasing demand in case of negative prices. It also influences the base load capacity (as it offers flexibility) and the transmission system (e.g., congestion).

approach to analyze strategic behavior in electricity markets[4]. TOU pricing can also reduce the inefficiency of single pricing, while being more practical, for most consumers, than the RTP of the wholesale market (see Celebi and Fuller (2007) and references therein for a review of TOU pricing).

Many electricity market models have either mostly ignored the demand response to changing prices (e.g., day-ahead models with mostly fixed demand), or, at the other extreme, they assumed that the full demand response occurred within one hour, see Day et al. (2002), Helman and Hobbs (2010), Hobbs (2001), Hobbs and Pang (2004). In the former case, it usually creates a problem especially in Cournot models, where the prices can rise infinitely, due to inelastic demand (Day et al., 2002). The proposed model of this thesis seeks to resolve this problem by bridging the speed of response gap between suppliers and consumers. The speed of response gap is the difference between the consumers' response to changes in price, which is no more frequent than the billing cycle allows (e.g., monthly), and the change in marginal costs of production, which occurs much more rapidly (e.g., hourly). In this thesis, we simulate consumer prices using monthly demand functions rather than hourly, with cross-price effects in the TOU models. TOU or single prices for consumers are determined for several months instead of each hour. Moreover, the demand response is dynamic in the model through a dependence of this month's demand on the previous month's demand. Note that these models are short-term (six months to a year) and the "lag" dependence on the previous month mostly represents the consumers' habits rather than long-term investment decisions (i.e., adjustment in capital stock is not considered).

The models proposed in this thesis address the problem of demand response with the inclusion of TOU pricing scheme in the context of bilateral markets, but the

[4] In actual electricity markets, power and its price are also differentiated by space (location on the transmission network) and fuel type (renewable versus fossil-fuelled or nuclear energy) and this may cause additional opportunities to exercise market power (see Hobbs and Pang (2004) and references therein). Note that a competitive rent earned by low-cost generators is not deemed to be evidence of market power.

models also apply to Ontario's POOLCO system because Ontario has a uniform pricing mechanism (i.e., no locational marginal prices)[5]. Hence, we have only considered the bilateral market case, where the role of the ISO is to coordinate the market participants (e.g., in a way that minimizes suppliers' operating costs) by using an incentive (e.g., penalty/payment) mechanism to curtail any trade that is not following the historical shape of the load duration curve[6].

In this thesis, our main focus remains in, but not limited to, the Ontario electricity market, since there is a move towards TOU pricing scheme with enhanced metering investment in Ontario. As a recent study by the Ontario Energy Board (OEB) (2007) showed, under TOU prices, consumers are able to reduce their overall consumption (conservation effect) and shift their consumption to mid-peak and off-peak periods (demand response effect). Other pilot projects are underway to fine tune TOU pricing schemes so that consumers are encouraged to control their consumption and cost. This move in Ontario towards TOU pricing with substantial investment in metering technology will certainly affect the market structure in the future. Within this context, the proposed model can be used as a forecasting and policy analysis tool (e.g., to assess potential market power of large suppliers) by regulatory bodies (e.g., Market Surveillance Panel of OEB and Independent Electricity System Operator –IESO).

There have been several studies in the literature about the efficiency and welfare effects of RTP and comparisons among single pricing (e.g., seasonal or monthly flat-

---

[5] In the presence of arbitrage (that erase any non-cost based differences in prices) and a network representation, Cournot competition in a bilateral market is equivalent to Cournot competition among generators in a POOLCO (i.e., generators sell to a central auction) (Hobbs, 2001). Without network representation, POOLCO and bilateral market models would be equivalent, too. Arbitrage is not relevant in this case, since there are no non-cost price differences (i.e., no locational marginal price differences).

[6] Load duration curve is obtained by re-ordering the hourly load (demand) data in descending order of magnitude, rather than chronologically for a period of time (e.g., year or month). The area under the load duration curve is the total electricity (e.g., kWh) used per period. It is most widely used for "resource planning" (e.g., medium- to long-term planning of power production) (Kirschen and Strbac, 2004 ; Conejo et al., 2006).

rates) and seasonal TOU (i.e., TOU prices that change between winter and summer months) (Borenstein, 2005; Borenstein and Holland, 2005; Holland and Mansur, 2006). These studies have estimated the size of potential efficiency gains from RTP adoption and have performed reasonable sensitivity analysis for their simulations. In particular, Borenstein and Holland (2005) have calculated the long-run efficiency gains (e.g., considering capacity investments and long-term welfare transfers) of 3% to 11% of the energy bill with RTP adoption and Borenstein (2005) has simulated about a quarter of these gains for seasonal TOU pricing. Holland and Mansur (2006) have found that the short-run efficiency gains (e.g., short-run welfare and consumers' surplus gains) are modest (0.24% and 2.5% of the total energy bill, respectively) if all customers adopt RTP. They have stated that monthly flat rates or seasonal TOU prices can only capture one third or a quarter of efficiency gains with RTP adoption, respectively. Moreover, they have analyzed the environmental effects of RTP adoption.

Our approach in this thesis is different in several ways. Aforementioned studies specify constant elasticity demand functions for each hour, whereas we propose monthly linear demand functions. They also assume zero cross-price elasticities among hours or time-blocks; while we use non-zero cross-price elasticities. Also, they do not explicitly model the behavioral aspects and dynamics (e.g., habit formation with lag effect) of the demand side. Furthermore, their analyses are in the context of perfect competition only, but they have stated that market power would increase the efficiency gains. In our illustrations in section 2.3.2, we have compared the monthly TOU prices over monthly single prices for different market structures.

Determining the degree of market power exercised by suppliers is a problem for regulators in electricity markets. Typical measures of market power include the market shares, pivotal supplier measure, concentration indices (e.g., Hirschmann-Herfindahl Index –HHI) and PCMs (e.g., Lerner index) (see Helman (2006) for discussion). This thesis shows how to conduct a welfare analysis by using an approximation method similar to Harberger (1971) to calculate the change in consumer surplus. A benchmark model of perfect competition is used to examine the surplus changes and PCMs for different market structures.

We represent our models using the VI framework, which is an effective and convenient way to create and manage our models. The VI formulations are developed to estimate *ex ante* TOU prices in different market structures, namely, perfect competition, oligopoly (i.e., all firms or several large firms compete à la Cournot[7]) and monopoly. The aim is to see the range of price manipulations for different structures. A scenario of the break-up of a large generation company into smaller parts is also examined (e.g., breaking-up the largest supplier in Ontario, Ontario Power Generation – OPG into two or more parts).

The supply side of the model is deliberately simplified here. The Ontario electricity market has a uniform (single) market clearing price system rather than a nodal/zonal pricing system (i.e., locational marginal prices), and hence transmission network representation is not applicable for the Ontario market. But a network representation must be added for nodal/zonal pricing systems (e.g., Pennsylvania New Jersey Maryland –PJM, New York, New England). However, the network representation would increase the size of the problem considerably and may require special algorithms to alleviate the complexity of the problem (e.g., decomposition algorithms as in Fuller and Chung (2005, 2008); also refer to chapter 3 and 4). Linearized DC network for TOU pricing models is introduced in chapter 3.

2.2 *Time of Use Pricing Models*

This subsection presents a multi-firm, multi-period equilibrium model in electricity markets with TOU pricing. The model consists of three parts: the ISO's problem, supply side (e.g., firm *f*'s problem) and the demand side. Symbols for the ISO

[7] Cournot oligopoly is the most common framework to model interaction among participants in electricity markets. In this framework, a supplier takes its rivals' sales and/or production quantities as fixed within its profit maximization problem. Other oligopolistic models (e.g., supply function, Bertrand, Stackelberg, tacit collusion) can be examined but the Cournot model is the most practical (Helman and Hobbs, 2010). Furthermore, it may be sufficient to simulate market prices (Sioshansi and Oren, 2007).

and supply side problems are defined in the following list. Symbols for the demand side are defined in section 2.2.4.

Sets

set of generation facilities: $i=1,\dots,I$

set of demand blocks: $j=1,\dots,J$ (alias index k)

set of periods (months): $t=1,\dots,T$

set of hours in a demand block j, period t: $h=1,\dots,H_j^{(t)}$ (defined by the market regulator)

set of firms: $f=1,\dots,F$

Parameters

$c_{fi}^{(t)}$ = operating cost per unit of energy for firm f's facility i in period t (\$/MWh)

$\kappa_{fi}^{(t)}$ = capacity of firm f's facility i in period t (MW)

$\delta_{jh}^{(t)}$ = fraction of total energy demand during block j of month t that occurs during hour h (see section 2.2.2 for explanation of this load shape parameter).

Decision variables

$z_{fijh}^{(t)}$ = the energy flowing from firm f's facility i to demand block j for hour h in period t (MWh)

$d_{fj}^{(t)}$ = sales by firm f to demand block j in period t (MWh)

$p_j^{(t)}$ = TOU prices (e.g., uniform block prices) in period t for demand block j (off-peak, mid-peak and on-peak, \$/MWh) (a function of $d_{fj}^{(t)}$ variables, but treated as a parameter by price-taking firms)

2.2.1 ISO's Problem

In this model, the role of the ISO is minimized (e.g., a "minISO") and the ISO has no information about the financial arrangements (e.g., price, terms and conditions) among the buyers and sellers (Varaiya and Wu, 1997; Wu and Varaiya, 1999; Bhattacharya et al., 2001; Shahidehpour, 2001; Stoft, 2002). The ISO only acts as a medium-term (e.g., for six months) planner for generation scheduling (rather than hourly dispatch). It minimizes the overall operating costs of the suppliers while inciting commitments by using the historical shape of the load duration curve. The operating procedures in this setting are as follows:

Suppliers (e.g., generator firms) inform ISO about their bilateral trade quantities (e.g., sales of each firm to demand block *j*, for month *t*, $d_{fj}^{(t)}$) a month ahead for each demand block *j*. In ISO's problem, the load variation within a demand block is enforced to follow the historical pattern (by using the parameters, $\delta_{jh}^{(t)}$, that are explained in section 2.2.2). Since the transmission network is not represented, the feasibility of trades (e.g., system balance and congestion management) is not a concern (but it is included in chapter 3).

- Suppliers and consumers (or consuming entities such as local distribution utilities) make bilateral trades and ISO is informed about the demand (load) schedules for each demand block *j* a month ahead.
- ISO enforces overall trades to follow the historical pattern of the load duration curve by broadcasting the penalty/payment scheme and hourly generation schedules.
- With this information, suppliers and consumers modify their trades accordingly and return new trades.
- This process can iterate between ISO and suppliers/consumers until equilibrium is reached. We can also assume that suppliers can anticipate the penalty/payment scheme (e.g., $\lambda_{jh}^{(t)}$, in (2.1)) for their profit maximization problem and this iterative process for equilibrium can be avoided.
- Finally, the generation and demand schedules for month *t*, demand block *j* and hour *h* are committed and all trades become firm (e.g., non-curtailable).

The ISO's penalty/payment scheme does not create any revenue or profit for the ISO, but rather, it acts as an incentive mechanism that shifts money around among firms (see section 2.3). In other words, the ISO encourages firms to minimize their deviations from the historical pattern of the load variation within a demand block *j*.

$$
\begin{aligned}
&\max_{z} \; -\sum_{f=1}^{F}\sum_{i=1}^{I}\sum_{h=1}^{H_j^{(t)}} c_{fi}^{(t)} z_{fijh}^{(t)} \\
&subject\ to && && [dual] \\
&d_{fj}^{*(t)} - \sum_{i=1}^{I}\sum_{h=1}^{H_j^{(t)}} z_{fijh}^{(t)} \le 0 && \forall\, f && \left[\rho_{fj}^{(t)}\right] \\
&z_{fijh}^{(t)} \le \kappa_{fi}^{(t)} && \forall\, f,i,h && \left[\mu_{fijh}^{(t)}\right] \\
&\delta_{jh}^{(t)} \sum_{f=1}^{F} d_{fj}^{*(t)} - \sum_{f=1}^{F}\sum_{i=1}^{I} z_{fijh}^{(t)} = 0 && \forall\, h && [\,\lambda_{jh}^{(t)}] \\
&z_{fijh}^{(t)} \ge 0 && \forall\, f,i,h
\end{aligned}
\tag{2.1}
$$

In (2.1), for each block of hours *j*, and for each month *t*, the ISO chooses $z_{fijh}^{(t)}$ to minimize total operating costs of all generation facilities of firms that inject amounts $d_{fj}^{(t)}$ during demand block *j*, while respecting the fluctuations in demand over the hours of blocks. Here, we assume that each supplier firm *f* reports its bilateral contract amount $d_{fj}^{(t)}$ (but not the prices $p_j^{(t)}$) to the ISO before each month *t*. Furthermore, it is assumed that the ISO is aware of the marginal cost (e.g., $c_{fi}^{(t)}$) and capacity of each generation facility (e.g., $\kappa_{fi}^{(t)}$) at each month *t*. In the ISO's problem, the amount of each firm's sales, $d_{fj}^{(t)}$, is treated as fixed (denoted by superscript *). The first set of constraints in (2.1) ensures that electricity supply of firm *f* is sufficient to meet its sales to demand block *j*. This constraint assumes that each bilateral trade is firm (e.g., non-curtailable) and firms have to satisfy their bilateral agreements. In other words, ISO requires each firm to satisfy their bilateral agreements, which is not a common practice in actual markets. However, for planning purposes, ISO may assume all bilateral trades will be realized. The second set of constraints contains the capacity constraint for each generation facility owned by firm *f*. The next subsection explains in detail how the third set of constraints ensures that generation matches demand in every hour of block *j*.

2.2.2 Estimation of Load Variation within a Demand Block

We now explain the meaning and measurement of the parameters $\delta_{jh}^{(t)}$ which link the different time scales of the supply and demand sides of the model.

When consumers pay the same price for energy at any hour within a TOU block, it is reasonable to suppose that the demand variations over hours within the block are related to non-price causes, such as temperatures, natural lighting, daily meal schedules and habits of all kinds that affect electricity usage. Such non-price causes must necessarily be represented by parameters, not variables to be solved for. Celebi and Fuller (2007) proposed to measure the pattern of variation in demand that has been observed in the recent past, and to assume that the same pattern (but not the absolute values) will repeat in the near future, i.e., within the model's time horizon. These parameters for a specific demand block j can be calculated from historical observations of a previous year, as the fraction, $\delta_{jh}^{(t)}$, of total energy demand in block j, month t, that occurs during hour h:

$$\delta_{jh}^{(t)} = d_{jh}^{(t)} / d_{j}^{(t)}$$

where $d_{jh}^{(t)}$ is the hourly energy demand for hour h within demand block j and $d_{j}^{(t)}$ is the total energy demand in block j (i.e., $d_{j}^{(t)} = \sum_{h=1}^{H_j^{(t)}} d_{jh}^{(t)}$). A property of these parameters is that summation over hours of demand block j is 1, i.e., $\sum_{h=1}^{H_j^{(t)}} \delta_{jh}^{(t)} = 1$.

The third set of constraints in (2.1) states that the hourly generation for all different facilities and all firms should meet the total sales of all firms at hour h. With this condition, the ISO imposes the historical shape of the load duration curve within the hours of demand block j. But if prices differ from historical ones, then the entire month's load duration curve of the solution can have a shape that is different from the historical shape. The dual variable $\lambda_{jh}^{(t)}$ (unconstrained in sign) for this condition is the penalty/payment for firm f's hourly deviations of its sales from the average hourly demand in demand block j and hourly deviations of its output $\sum_{i=1}^{I} z_{fijh}^{(t)}$ from the average output over all hours in the block j. (i.e., deviations from the historical shape of the load duration curve).

2.2.3 Supply Side: Firm f's Problem

The supply side of the model, formulated in (2.2), maximizes firm f's profit, π_f, i.e., the total revenues of firm f minus the total operating cost of firm f's hourly generation by different technologies of production (e.g., nuclear, hydro, coal, gas/oil, indexed by i) to meet its sales in different demand blocks (e.g., off-peak, mid-peak, on-peak, indexed by j) plus the ISO's penalty/payment due to variations from hourly average demand and hourly average output in demand block j. Here, we assume that supplier firms are able to anticipate ISO's penalty/payment scheme (e.g., $\lambda_{jh}^{*(t)}$).

$$
\begin{aligned}
&\max_{d,z} \pi_f = \sum_{t=1}^{T}\sum_{j=1}^{J}\left[p_j^{(t)} - \sum_{h=1}^{H_j^{(t)}} \delta_{jh}^{(t)} \lambda_{jh}^{*(t)}\right] d_{fj}^{(t)} - \sum_{t=1}^{T}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{h=1}^{H_j^{(t)}}\left[c_{fi}^{(t)} - \lambda_{jh}^{*(t)}\right] z_{fijh}^{(t)} \\
&\textit{subject to} \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad [dual] \\
&d_{fj}^{(t)} - \sum_{i=1}^{I}\sum_{h=1}^{H_j^{(t)}} z_{fijh}^{(t)} \le 0 \qquad \forall\, j,t \qquad \left[\rho_{fj}^{(t)}\right] \qquad (2.2) \\
&z_{fijh}^{(t)} \le \kappa_{fi}^{(t)} \qquad \forall\, i,j,h,t \qquad \left[\mu_{fijh}^{(t)}\right] \\
&z_{fijh}^{(t)} \ge 0 \qquad \forall\, i,j,h,t
\end{aligned}
$$

Various market structures are modeled within this framework. The perfect competition structure –with firms treating $p_j^{(t)}$ as a parameter beyond their control– serves as a reference case, as it would lead to the most efficient market performance. On the other hand, the monopoly structure represents the worst outcome of exercising market power. In between is the Nash-Cournot structure where either all firms or some large firms act à la Cournot. In the monopoly and Nash-Cournot structures, firms see their knowledge of the dependence of $p_j^{(t)}$ on total market demand, as detailed in section 2.3.

As explained in Celebi and Fuller (2007), the supply model can be extended to be more realistic as long as each firm's model remains as a LP (e.g., a linearized DC network, line limits and ramping constraints can be included at the expense of problem size) or more generally as a convex program.

In firm f's problem, it is presumed that the ISO's penalty/payment term, $\lambda_{jh}^{(t)}$, is fixed (denoted by superscript *). There is no discounting because the model is intended to cover only a few months into the future; discounting could be included for longer time horizons. The first set of constraints ensures that electricity supply of firm f is sufficient to meet its sales to demand block j; at an optimal solution, these constraints are binding equalities. The second set of constraints contains the capacity constraint for each generation facility owned by firm f. It should be noted that the ISO's and firm f's problem have the common variable, $z_{fijh}^{(t)}$, which, in equilibrium, are equivalent as shown in section 2.3.

2.2.4 Demand Side

The demand side is represented by demand equations that use the prices and lagged demand as independent variables, such that the reaction of demand to changes in price is a process in time (Celebi and Fuller, 2007; Taylor and Houthakker, 2010). Especially in energy markets, the adjustment to varying prices can occur after some periods rather than instantaneously. The response of the consumer is not at a point of time but rather distributed over time because of usage patterns (i.e., habits), imperfect information about the market and need for some uninterruptible services that are supplied by energy used by stocks of equipment –e.g., refrigerators, lights. Pollak (1970; 1990) formulated a model of consumer behavior based on "habit formation" (also see Taylor and Houthakker (2010) for dynamic demand models). His "habit formation" assumptions imply that consumption in the previous period (month) influences current preference and demand. A fundamental assumption of the habit formation model is that the individual consumer does not take account of the effect of his current purchase on his future preferences and future consumption. In the case of habit formation, this assumption is plausible; in the case of capital stock (e.g., consumer durables) adjustment, it is not. For our demand model, we have used the former model rather than the latter, as capital stock adjustment would be more appropriate for long-term models. A model of demand for this adjustment must explicitly recognize the intertemporal nature of the problem. We have focused on the short-term analysis of demand behavior and the effect of capital stock adjustment is not considered. In our models, we also assume that consumers' decision making process is sequential in time and separate for each period. Hence, lagged demands are treated as parameters in the utility maximization problem of the individual consumer at each period (see Appendix A for details). A distributed lag model can represent this response process in time. One form of a one commodity model is the linear distributed lag model:

$$d^{(t)} = a^{(t)} + b^{(t)}p^{(t)} + e^{(t)}d^{(t-1)} \tag{2.3}$$

where $d^{(t)}$ is the demand of electricity in period t, $a^{(t)}$ is a constant representing non-price effects (e.g., weather conditions, socio-demographic factors), $p^{(t)}$ is the price of electricity at period t, $d^{(t-1)}$ is the lagged demand, $b^{(t)}$ is the price coefficient at period t and $e^{(t)}$ is the lag coefficient at period t. In a real world application, a careful econometric study would be needed, to establish the best functional form, and its parameters.

We allow the parameters $b^{(t)}$ and $e^{(t)}$ to vary with time because, in our illustration of section 2.3.1, we estimate these parameters by a linear approximation of the constant elasticity demand model around the perfect competition solution of Celebi and Fuller (2007). For example, in (2.3), $b^{(t)} = \varepsilon\, \bar{q}^{(t)}/\bar{p}^{(t)}$, where ε is the constant elasticity (independent of time), and $\bar{q}^{(t)}$ and $\bar{p}^{(t)}$ are the quantity and price at time t; $b^{(t)}$ depends on time, unless the ratio $\bar{q}^{(t)}/\bar{p}^{(t)}$ happens to be constant in time.

Equation (2.3) can be extended to a multi-commodity case where each commodity is the electricity demand in different times of day (e.g., demand blocks: on-peak, mid-peak, off-peak):

$$\boldsymbol{D}^{(t)} = \boldsymbol{A}^{(t)} + \boldsymbol{B}^{(t)}\boldsymbol{P}^{(t)} + \boldsymbol{E}^{(t)}\boldsymbol{D}^{(t-1)} \tag{2.4}$$

where

$\boldsymbol{A}^{(t)}$ = vector of the factors representing non-price effects at period t,

$\boldsymbol{D}^{(t)}$ = vector of all demands for electricity in period t (i.e., on-peak, mid-peak, off-peak demand) where $\boldsymbol{D}^{(t)} = \left[d_j^{(t)}\right]$ and $d_j^{(t)} = \sum_{f=1}^{F} d_{fj}^{(t)}$.

$\boldsymbol{P}^{(t)}$ = vector of TOU electricity prices in period t (i.e., on-peak, mid-peak, off-peak prices)

$\boldsymbol{B}^{(t)}$ = a square matrix of the price coefficients (i.e., own-price and cross-price) for period t

$\boldsymbol{E}^{(t)}$ = a square diagonal matrix of the lag coefficients for period t.

$\boldsymbol{B}^{(t)}$ is assumed to be invertible for our analyses (i.e., inverse demand functions are well defined). Moreover, if the lagged demand terms in (2.4) are treated as parameters at each period (i.e., consumers' decision making process is separate and sequential for each period), consumers' preferences and their underlying utility maximization problem can be recovered. This recovery of preferences is the basis for the consumers' welfare analysis (see footnote 28 in Appendix A for these recovery conditions). Many studies have shown how to approximate the change in consumers' surplus using prices and quantities for different scenarios (see econometric studies specifically on TOU pricing in Acton and Mitchell (1980) and Caves et al. (1984)). We have shown in the Appendix A that the change in consumers' surplus that accounts for the effect of lagged demand can be calculated using prices and quantities of different scenarios similar to Harberger's approximation[8] (Harberger, 1971). Also in the short-

[8] If the model has symmetric demand, and therefore an objective to maximize welfare (normally consumers' plus producers' surplus), a policy analyst can observe the change in surplus between two different runs of the equilibrium model. When the demand functions are not symmetric (non-integrable) an approximation method which was introduced by Arnold Harberger (1971) can be used to estimate the change in consumers' surplus. He used a Taylor series approximation for the change in total value for a single consumer. Change in consumers' surplus is then given by the change in consumers' total value minus change in consumers' payments, summed over all commodities j in the model:

$\sum_{j=1}^{n}\left(p_j + \frac{1}{2}\Delta p_j\right)\Delta X_j - \left[(p_j + \Delta p_j)(X_j + \Delta X_j) - p_j X_j\right]$,

where p_j is the previous equilibrium price, X_j is the quantity demanded in the previous equilibrium, Δp_j is the change in prices and, Δx_j is the change in quantities demanded (Celebi and Fuller, 2007).

term, we have assumed that marginal utility of income is constant (i.e., effect of expenditure on energy commodities is not very significant in the budget). This ensures that the change in consumer surplus is a meaningful money measure of utility change in case of multiple price changes (i.e., path independence of price adjustments) (Just et al., 2004).

2.3 *Analyses of Different Market Structures*

The models proposed in this thesis are represented and solved by the VI problem approach. To aid readers who are familiar with MCP models, but not VI problems, we first formulate the perfect competition model as a MCP, followed by the VI form and a justification for the equivalence of the two forms. The oligopoly and monopoly models are presented only in the VI form.

In (2.5), we formulate the perfect competition model as a MCP, by writing out the necessary Karush-Kuhn-Tucker (KKT) conditions for the ISO's and firm f's problems along with the demand equation, where $a_j^{(t)}$and $p_j^{(t)}$ are the jth elements of vectors $\boldsymbol{A}^{(t)}$and $\boldsymbol{P}^{(t)}$, respectively. Similarly, $b_{jk}^{(t)}$ and $e_{jj}^{(t)}$ are the elements of matrices $\boldsymbol{B}^{(t)}$ and $\boldsymbol{E}^{(t)}$, respectively.

MCP: Find $d_{fj}^{(t)}, z_{fijh}^{(t)}, p_j^{(t)}, \rho_{fj}^{(t)}, \mu_{fijh}^{(t)}, \lambda_{jh}^{(t)}$ that satisfy

$$
\begin{aligned}
&d_{fj}^{(t)} \geq 0 \perp && -p_j^{(t)} + \sum_{h=1}^{H_j^{(t)}} \delta_{jh}^{(t)} \lambda_{jh}^{(t)} + \rho_{fj}^{(t)} \geq 0 && \forall f, j, t \\
&z_{fijh}^{(t)} \geq 0 \perp && c_{fi}^{(t)} - \rho_{fj}^{(t)} + \mu_{fijh}^{(t)} - \lambda_{jh}^{(t)} \geq 0 && \forall f, i, j, h, t \\
&\rho_{fj}^{(t)} \geq 0 \perp && d_{fj}^{(t)} - \sum_{i=1}^{I} \sum_{h=1}^{H_j^{(t)}} z_{fijh}^{(t)} \leq 0 && \forall f, j, t \\
&\mu_{fijh}^{(t)} \geq 0 \perp && z_{fijh}^{(t)} \leq \kappa_{fi}^{(t)} && \forall f, i, j, h, t \\
&\lambda_{jh}^{(t)} free \perp && \delta_{jh}^{(t)} \sum_{f=1}^{F} d_{fj}^{(t)} - \sum_{f=1}^{F} \sum_{i=1}^{I} z_{fijh}^{(t)} = 0 && \forall j, h, t \\
&\sum_{f=1}^{F} d_{fj}^{(t)} = a_j^{(t)} + \sum_{k=1}^{K} b_{jk}^{(t)} p_k^{(t)} + e_{jj}^{(t)} \sum_{f=1}^{F} d_{fj}^{(t-1)} && && \forall j, t
\end{aligned}
\tag{2.5}
$$

The first four conditions in (2.5) are the necessary KKT conditions for firm *f*'s problem. The second to fifth conditions in (2.5) are the necessary KKT conditions for the ISO's problem (i.e., second to fourth conditions are common for the ISO's and firm *f*'s problem). The last equation is the linear distributed lagged demand equation. Note that the first condition does not include the extra term for the marginal revenue of firm f (i.e., recognizing that $p_j^{(t)}$ is a function of $\sum_{f=1}^{F} d_{fj}^{(t)}$) that appears in monopoly and Cournot models. Because all firms are price takers in a perfect competition structure, this condition only has the $p_j^{(t)}$ term as the marginal revenue term. Also note that the third and fifth conditions in (2.5) are linearly dependent at a solution, where the third constraints are binding (i.e., summing the third condition over all firms f equals the sum of the fifth condition over all hours h) and one combination of t, f, j in the third set of constraints can be dropped from (2.5).

When $d_{fj}^{(t)} > 0$ (implying $z_{fijh}^{(t)} > 0$ from the third condition), the first two conditions in (2.5) become equalities and we can derive the following condition by summing them:

$$p_j^{(t)} - \sum_{h=1}^{H_j^{(t)}} \delta_{jh}^{(t)} \lambda_{jh}^{(t)} + \lambda_{jh}^{(t)} = c_{fi}^{(t)} + \mu_{fijh}^{(t)} \tag{2.6}$$

The left hand side can be understood as the hourly price; let it be denoted by $p_{jh}^{(t)}$. All firms receive this hourly price whereas consumers are paying the TOU price ($p_j^{(t)}$) in the models. Note that firms have adequate revenue when $z_{fijh}^{(t)} > 0$, because $p_{jh}^{(t)} - c_{fi}^{(t)} = \mu_{fijh}^{(t)} \geq 0$. If we multiply the left hand side of (2.6) by $\delta_{jh}^{(t)}$ and sum over all hours h, we derive the condition, $p_j^{(t)} = \sum_{h=1}^{H_j^{(t)}} \delta_{jh}^{(t)} p_{jh}^{(t)}$. This is the weighted average condition imposed in Celebi and Fuller (2007) (without the discount factor), which relates the hourly marginal cost $p_{jh}^{(t)}$ to consumer TOU price $p_j^{(t)}$ in the perfect competition case. This also ensures that the revenue requirement of all firms for demand block j is met by revenue collected from consumers.

$$p_j^{(t)} \sum_{f=1}^{F} d_{fj}^{(t)} = \sum_{h=1}^{H_j^{(t)}} \delta_{jh}^{(t)} p_{jh}^{(t)} \sum_{f=1}^{F} d_{fj}^{(t)} = \sum_{h=1}^{H_j^{(t)}} \sum_{f=1}^{F} \sum_{i=1}^{I} p_{jh}^{(t)} z_{fijh}^{(t)} \tag{2.7}$$

Related to this is the fact that the penalties/payments imposed by the ISO sum to zero, over all firms, within every block *j*:

$$\begin{aligned} &-\sum_{f=1}^{F} \sum_{h=1}^{H_j^{(t)}} \delta_{jh}^{(t)} \lambda_{jh}^{(t)} d_{fj}^{(t)} + \sum_{f=1}^{F} \sum_{i=1}^{I} \sum_{h=1}^{H_j^{(t)}} \lambda_{jh}^{(t)} z_{fijh}^{(t)} \\ &= -\sum_{h=1}^{H_j^{(t)}} \sum_{f=1}^{F} \sum_{i=1}^{I} \lambda_{jh}^{(t)} z_{fijh}^{(t)} + \sum_{f=1}^{F} \sum_{i=1}^{I} \sum_{h=1}^{H_j^{(t)}} \lambda_{jh}^{(t)} z_{fijh}^{(t)} = 0 \end{aligned} \tag{2.8}$$

Thus the ISO's penalty/payment scheme shifts money around among firms, but does not directly involve consumers.

We can also formulate (2.5) as a VI problem. The feasible set for the VI problem is defined as follows:

$$K = \left\{ d_{fj}^{(t)}, z_{fijh}^{(t)}, p_j^{(t)} \left| \begin{array}{ll} d_{fj}^{(t)} - \sum_{i=1}^{I} \sum_{h=1}^{H_j^{(t)}} z_{fijh}^{(t)} \leq 0 & \forall\, f, j, t \\ z_{fijh}^{(t)} \leq \kappa_{fi}^{(t)} & \forall\, f, i, j, h, t \\ z_{fijh}^{(t)} \geq 0 & \forall\, f, i, j, h, t \\ \delta_{jh}^{(t)} \sum_{f=1}^{F} d_{fj}^{(t)} - \sum_{f=1}^{F} \sum_{i=1}^{I} z_{fijh}^{(t)} = 0 & \forall\, j, h, t \\ \sum_{f=1}^{F} d_{fj}^{(t)} = a_j^{(t)} + \sum_{k=1}^{K} b_{jk}^{(t)} p_k^{(t)} + e_{jj}^{(t)} \sum_{f=1}^{F} d_{fj}^{(t-1)} & \forall\, j, t \end{array} \right. \right\}$$

Note that $p_j^{(t)}$ variables are implicitly defined by the $d_{fj}^{(t)}$ variables. Instead, an explicit inverse demand function can be used for a more compact formulation without $p_j^{(t)}$ variables, but for ease of readability of the formulation, the $p_j^{(t)}$ variables are used.

In the feasible set K, the first four constraints are from the ISO's and firm *f*'s problems and the last equation is the linear distributed lagged demand equation.

The VI problem for the perfect competition model is as in (2.9). To relate (2.9) to the general VI form (1.1), the vector $\boldsymbol{x}$ contains the variables $d_{fj}^{(t)}, z_{fijh}^{(t)}$ and $p_j^{(t)}$ for all

f, i, j, h and t, and the elements of the vector-valued mapping $\boldsymbol{F}(\boldsymbol{x})$ are as follows: $-p_j^{(t)}$ is the element of $\boldsymbol{F}$ that corresponds to $d_{fj}^{(t)}$; $c_{fi}^{(t)}$ is the element of $\boldsymbol{F}$ that corresponds to $z_{fijh}^{(t)}$; and the element of $\boldsymbol{F}$ that corresponds to $p_j^{(t)}$ is zero. Note that for $d_{fj}^{(t)}$ and $z_{fijh}^{(t)}$, the corresponding elements of $\boldsymbol{F}$ are the partial derivatives of the objective function of firm f's problem (2.2) (i.e., the $\lambda_{jh}^{(t)}$ terms are cancelled out.)

Find $\left(d_{fj}^{(t)*}, z_{fijh}^{(t)*}, p_j^{(t)*}\right) \in K$ such that

$$-\sum_{t=1}^{T}\sum_{f=1}^{F}\sum_{j=1}^{J} p_j^{(t)*}\left(d_{fj}^{(t)} - d_{fj}^{(t)*}\right) + \sum_{t=1}^{T}\sum_{f=1}^{F}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{h=1}^{H_j^{(t)}} c_{fi}^{(t)}\left(z_{fijh}^{(t)} - z_{fijh}^{(t)*}\right) \geq 0 \quad \forall \left(d_{fj}^{(t)}, z_{fijh}^{(t)}, p_j^{(t)}\right) \in K \tag{2.9}$$

The VI problem (2.9) in primal variables has the KKT conditions listed in (2.5) and hence is equivalent to the MCP (2.5) (see Harker and Pang (1990) and Nagurney (1993) for discussion of the KKT conditions for VI problems).

There is a minor technicality in the derivation of (2.5) from the KKT conditions of (2.9). Let $\upsilon_j^{(t)}$ be the dual variable of the distributed lagged demand equation in the definition of K, and let $\boldsymbol{\upsilon}^{(t)}$ be the vector with elements $\upsilon_j^{(t)}$. The KKT conditions which correspond to the $p_j^{(t)}$ variables, in vector-matrix form, are $\boldsymbol{B}^{(t)^T}\boldsymbol{\upsilon}^{(t)} = 0$, where $\boldsymbol{B}^{(t)^T}$ is the transpose of $\boldsymbol{B}^{(t)}$, the matrix of coefficients $b_{jk}^{(t)}$. Because $\boldsymbol{B}^{(t)}$ is assumed to be invertible (so that we can have inverse demand functions) it follows that $\boldsymbol{\upsilon}^{(t)} = 0$ for all t. Therefore, $\boldsymbol{\upsilon}^{(t)}$ and $\boldsymbol{B}^{(t)^T}\boldsymbol{\upsilon}^{(t)} = 0$ can be dropped from the list of KKT conditions of the VI problem, giving rise to the MCP (2.5).

We only provide the VI formulations for other market structures, for ease of representation. It is straightforward to derive their equivalent MCP formulations as in the perfect competition case. The other market structures have the same feasible set K. The VI problem for the Nash-Cournot model is formulated as follows:

Find $\left(d_{fj}^{(t)*}, z_{fijh}^{(t)*}, p_j^{(t)*}\right) \in K$ such that

$$-\sum_{t=1}^{T}\sum_{f=1}^{F}\sum_{j=1}^{J}\left(p_j^{(t)*} + \theta_{fj}^{(t)^*}\right)\left(d_{fj}^{(t)} - d_{fj}^{(t)*}\right) + \sum_{t=1}^{T}\sum_{f=1}^{F}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{h=1}^{H_j^{(t)}} c_{fi}^{(t)}\left(z_{fijh}^{(t)} - z_{fijh}^{(t)*}\right) \geq 0 \quad \forall\left(d_{fj}^{(t)}, z_{fijh}^{(t)}, p_j^{(t)}\right) \in K \tag{2.10}$$

where the term $p_j^{(t)^*} + \theta_{fj}^{(t)^*}$ is the marginal revenue for firm f in period t and demand block j, and $\theta_{fj}^{(t)^*}$ is the "extra" marginal revenue term. This marginal revenue term is derived from the partial derivative of the objective function in (2.2), with respect to $d_{fj}^{(t)}$, when the firm is aware of the price-quantity relation of the distributed lagged demand equation:

$$\frac{\partial \pi_f}{\partial d_{fj}^{(t)}} = p_j^{(t)} + \frac{\partial p_j^{(t)}}{\partial d_{fj}^{(t)}} d_{fj}^{(t)} + \frac{\partial p_j^{(t+1)}}{\partial d_{fj}^{(t)}} d_{fj}^{(t+1)} = p_j^{(t)} + \theta_{fj}^{(t)} \tag{2.11}$$

where $\left[\frac{\partial p_j^{(t)}}{\partial d_{fj}^{(t)}}\right] = \left(\boldsymbol{B}^{(t)}\right)^{-1}$ and $\left[\frac{\partial p_j^{(t+1)}}{\partial d_{fj}^{(t)}}\right] = -\boldsymbol{E}^{(t+1)}\left(\boldsymbol{B}^{(t+1)}\right)^{-1}$.

Note that the penalties/payments imposed by the ISO are neither included in the VI formulation (2.10) nor in the marginal revenue term, because they sum to zero, over all firms, within every block j.

As an extension to (2.10), we can also formulate a problem where several large firms act à la Cournot and the rest of the firms are price takers. This problem is formulated as follows:

Find $\left(d_{fj}^{(t)*}, z_{fijh}^{(t)*}, p_j^{(t)*}\right) \in K$ such that

$$-\sum_{t=1}^{T}\sum_{f=1}^{F}\sum_{j=1}^{J}\left(p_j^{(t)*} + \underset{f\neq g}{\theta_{fj}^{(t)}}{}^{*}\right)\left(d_{fj}^{(t)} - d_{fj}^{(t)*}\right) + \sum_{t=1}^{T}\sum_{f=1}^{F}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{h=1}^{H_j^{(t)}} c_{fi}^{(t)}\left(z_{fijh}^{(t)} - z_{fijh}^{(t)*}\right) \geq 0 \quad \forall\left(d_{fj}^{(t)}, z_{fijh}^{(t)}, p_j^{(t)}\right) \in K \tag{2.12}$$

where g is the index of the set of price taker firms (i.e., they don't have the "extra" marginal revenue term ${\theta_{fj}^{(t)}}^*$) .

Lastly we can define a VI problem for the monopoly structure:

Find $\left(d_{fj}^{(t)*}, z_{fijh}^{(t)*}, p_j^{(t)*}\right) \in K$ such that

$$-\sum_{t=1}^{T}\sum_{f=1}^{F}\sum_{j=1}^{J}\left(p_j^{(t)*} + \sum_{f'=1}^{F}\theta_{f'j}^{(t)}\right)\left(d_{fj}^{(t)} - d_{fj}^{(t)*}\right) + \sum_{t=1}^{T}\sum_{f=1}^{F}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{h=1}^{H_j^{(t)}} c_{fi}^{(t)}\left(z_{fijh}^{(t)} - z_{fijh}^{(t)*}\right) \geq 0 \quad \forall\left(d_{fj}^{(t)}, z_{fijh}^{(t)}, p_j^{(t)}\right) \in K \tag{2.13}$$

Note that f' is an alias index for f, and that all firms are owned by the monopolist.

For each VI problem (2.9), (2.10) and (2.12), the expression in the inequality is an estimate of the change in the negative of profits (summed over all firms) due to feasible deviations from equilibrium, using marginal revenues of the firms as measured at equilibrium. Therefore, at equilibrium, no firm sees any advantage in changing its variables $d_{fj}^{(t)}$ and $z_{fijh}^{(t)}$ in a feasible way. For the monopoly model (2.13), the monopolist firm sees no advantage in deviating from the equilibrium solution, i.e., the solution is a local maximum of profit (and a global maximum, due to convexity).

Instead of TOU pricing, some consumers may prefer a single price, or regulatory bodies in electricity markets may choose to implement a single pricing scheme. In this case, consumers' prices do not vary by time of day, and they may or may not vary by month (Ontario is now changing from single to TOU pricing). We illustrate by showing how to model a single price that is the same at all times of the day in a month. We can add a new constraint set to the feasible set K:

$$p_j^{(t)} = p_k^{(t)} \quad \forall\, j, k, t \text{ and where } j \neq k \tag{2.14}$$

With this set of constraints, the single prices may vary over different months. This is equivalent to having a single demand block (e.g., j with one element only). The single pricing model is solved with this additional constraint in the feasible set K using the same data and parameters for the TOU pricing model.

A careful reader would note the ease and convenience of representation of different market structures with VI formulations (compared to MCP formulations) and less effort required in coding these models in GAMS/EMP framework, because the code for the VI model does not require explicit representation of any dual variable. Both MCP and VI formulations in GAMS are available upon request from the authors.

2.3.1 An Illustrative Example: Assumptions and Data

We illustrate by using a past period's demand data for Ontario and find the TOU prices (*ex post*) for the same period. In real use, parameters carefully estimated from historical data would be used to forecast future TOU prices (*ex ante*). Moreover, in the regulatory context, sensitivity analysis may be required to demonstrate the robustness of the demand function specifications. This example consists of six periods (months May to October), four types of generation facilities (nuclear, hydro, coal, gas/oil), and three demand blocks (on-peak, mid-peak and off-peak electricity)[9]. There are only two firms (i.e., a duopoly for Nash-Cournot structure (2.10)) for this illustrative example. Firm 1 has 70% of the available generation capacity (almost equivalent to the total capacity of OPG) and firm 2 has the remaining 30% of the capacity (IESO, 2009). For another model, we assume that the firm 2 is a price taker in the Nash Cournot structure (2.12) with firm 1 acting à la Cournot, and we also solve the perfect competition (2.9) and monopoly (2.13) models.

The data for the models are obtained from Celebi and Fuller (2007), (except the generation capacities which are from IESO (2005)). The models have been reduced in size by using a representative weekday of the month instead of all days of the month. In Celebi and Fuller (2007), they have used a logarithmic demand equation. Here, we have

---

[9] In order to eliminate peculiarities in the first and last periods, due to lag and lead effects (resulting from the demand function and the marginal revenue term, respectively), we have used periods April to November in our actual analysis but reported the results for periods May to October only.

linearized their findings around the TOU pricing solution of the perfect competition model of a representative weekday of the month to obtain price ($\boldsymbol{B}^{(t)}$) and lag ($\boldsymbol{E}^{(t)}$) coefficient matrices. In other words, we have calibrated the demand parameters in the perfect competition model of Celebi and Fuller (2007) to obtain the same results in perfect competition model (2.9).

A major assumption is about the network structure. Transmission constraints such as line and voltage limits are ignored in our analyses. This means that there is a single price at any given time, as is now the case in Ontario. Geographically differentiated prices (i.e., nodal, zonal pricing) would require a representation of the transmission network in the model and this is introduced in chapter 3.

2.3.2 Results and Welfare Analysis

The models are coded in the GAMS/EMP framework and solved by the PATH solver[10]. Table 1 summarizes the TOU and single prices for different market structures and periods.

The perfect competition and monopoly models represent the two extreme market structures. The perfect competition model provides the lowest TOU prices for different blocks. At the other extreme, the monopoly model gives the highest TOU prices. The Nash-Cournot models with firm 1 as the only Nash-Cournot player or all firms as Nash-Cournot players are in between these two extreme market structures. In almost all demand blocks and time periods (except for on-peak, periods May, Sept. and Oct., and mid-peak, period Sept.), the TOU prices of the Nash-Cournot model with firm 1 as the only Cournot player are lower than the TOU prices of the Nash-Cournot model with all

---

[10] Most models are solved in less than a second on a 2.6 GHz Windows 2003 Server with 32 GB memory and 8 CPUs. In MCP and VI models with 2-firms, there are around 3,500 variables and conditions. Note that there are auxiliary variables (i.e., $p_j^{(t)}$) in both models and the model size can be reduced slightly (as much as $T \times J$) by eliminating them.

players acting à la Cournot. Note that the on-peak prices are highest in the periods June to August when the peak summer demand is expected.

Table 1: TOU and single prices ($/MWh) for different market structures

	Prices \ Period	May	June	July	Aug.	Sept.	Oct.
Perfect Competition	*off-peak*	32.36	32.47	33.80	36.31	31.53	33.57
	mid-peak	37.46	43.72	43.99	45.61	38.55	34.51
	on-peak	40.84	49.99	49.50	50.65	38.58	35.96
	single	**39.56**	**42.19**	**42.70**	**44.16**	**38.56**	**35.50**
Nash Cournot	*off-peak*	47.64	48.29	49.30	51.24	45.75	50.20
	mid-peak	55.47	63.74	63.67	66.49	56.19	52.15
	on-peak	59.18	71.53	71.00	72.55	55.75	56.13
	single	**56.94**	**60.79**	**62.06**	**65.55**	**55.36**	**53.18**
Nash Cournot (Firm 1 only)	*off-peak*	44.32	46.18	47.36	50.67	43.23	46.05
	mid-peak	54.49	60.12	60.36	61.40	56.63	51.86
	on-peak	59.97	66.68	67.18	68.33	56.75	56.36
	single	**53.16**	**57.08**	**58.56**	**61.72**	**51.96**	**51.67**
Monopoly	*off-peak*	56.01	56.32	58.28	63.24	54.90	60.49
	mid-peak	66.37	75.88	76.11	80.21	70.46	64.94
	on-peak	76.51	90.23	89.24	91.31	68.98	70.57
	single	**67.26**	**72.61**	**73.55**	**77.62**	**65.32**	**65.59**

The single prices for different months are the lowest for the perfect competition and the highest for the monopoly case, as expected. Similar to TOU prices, as the number of Cournot players increases, the single prices for different months increase.

We compute the PCMs, defined as the difference between a market structure's price and perfect competition price divided by the market structure's price (Bompard et al., 2005; Helman and Hobbs, 2010). PCMs (averaged over all periods) are presented in Table 2. Note that the Nash-Cournot structure with firm 1 as the single Nash-Cournot player has the lowest margins, and the monopoly has the highest.

Table 2: PCMs for TOU and single pricing models (averaged over all periods)

Market Structure \ PCMs	*Off-peak*	*Mid-peak*	*On-peak*	*Single*
Nash Cournot	31.6%	31.9%	31.4%	31.4%
Nash Cournot (Firm 1 only)	28.0%	29.5%	29.6%	27.4%
Monopoly	42.7%	43.9%	45.6%	42.5%

Table 3: Change in total of consumers', suppliers' and total surpluses over all periods (in thousand dollars) due to change from single pricing to TOU pricing

Market Structure \ Change in	*Consumers' Surplus*	*Suppliers' Surplus*	*Total Surplus*
Perfect Competition	6,611	-1,919	4,692
Nash Cournot	6,964	2,336	9,299
Nash Cournot (Firm 1 only)	2,795	3,862	6,657
Monopoly	275	4,574	4,850

Table 3 summarizes the welfare analyses (i.e., change in consumers', producers' and total surpluses) after the implementation of TOU prices over single pricing for different market structures. Total surplus and consumers' surplus increase for all market structures and suppliers' surplus increases for all structures except perfect competition.

In Table 4, we compare the perfect competition TOU model (i.e., reference case) to the other market structures. It is noted that the change in consumers' or total surplus is positive when the perfect competition model is compared to all other market structures.

Table 4: Change in total of consumers', suppliers' and total surpluses over all periods for TOU pricing due to change from oligopoly/monopoly to perfect competition (in thousand dollars)

Market Structure \ Change in	*Consumers' Surplus*	*Suppliers' Surplus*	*Total Surplus*
Nash Cournot	86,147	-36,736	49,411
Nash Cournot (Firm 1 only)	76,693	-30,496	46,197
Monopoly	134,777	-45,882	88,895

In Table 5, we compute the consumer surplus gains as percentage of the total energy bill. We have found modest consumer surplus gains of monthly TOU prices over monthly single pricing (e.g., monthly flat-rates) under different market structures. However, market power (changing market structure from oligopoly/monopoly to perfect competition under TOU pricing scheme) has some significant effect on these gains in our illustrations.

Table 5: Change in total of consumers' surpluses over all periods as percent of total energy bill

Change from / Market Structure	*Single to TOU Pricing*	*Oligopoly/Monopoly to Perfect Competition for TOU Pricing*
Perfect Competition	5.78%	N/A
Nash Cournot	5.20%	62.36%
Nash Cournot (Firm 1 only)	2.08%	56.25%
Monopoly	0.30%	94.77%

These comparisons can be extended (but not illustrated in this thesis for brevity) by modifying the constraints (2.14) for seasonal TOU pricing as:

$$p_j^{(t)} = p_j^{(t')} \qquad \forall\, j, t, t' \text{ and where } t \neq t'; \tag{2.15}$$

and for seasonal single pricing as:

$$p_j^{(t)} = p_k^{(t')} \qquad \forall\, j, k, t, t' \text{ and where } j \neq k \text{ and } t \neq t'. \tag{2.16}$$

This illustrative example suggests a strong support for TOU pricing over the single pricing scheme in different market structures. These welfare analyses can be used by regulatory bodies in determining whether to pursue TOU prices or single prices. The welfare gains from TOU prices can be compared with the investment in metering technology and communication infrastructure. However, plausible sensitivity analyses for the model parameters are required for any policy recommendations.

2.3.3 Break-up of the Large Firm into Two or More Parts

In this subsection we discuss the break-up of the large supplier (firm 1 or namely, OPG in our illustration) into two equal parts (i.e., two firms with each having 35% of the total available capacity) for a model with three Cournot firms. Then we compare it with the 2-firm Nash-Cournot (i.e., duopoly) structure. Here, we assume that the cost structures of the two new parts do not change with the break-up of the large firm.

In Table 6, we compare the 2-firm structure with the 3-firm structure (for TOU and single pricing schemes separately). The consumers' and total surpluses increase in both cases.

Table 6: Comparison of 2-firm and 3-firm oligopolies (in thousand dollars)

Change in / Market Structure	*Consumers' Surplus*	*Suppliers' Surplus*	*Total Surplus*
2-firm to 3-firm oligopoly (TOU pricing)	18,309	-5,833	12,477
2-firm to 3-firm oligopoly (Single pricing)	46,419	-14,225	32,194

For this illustrative example, we further break up the large firms and increase the number of Cournot players to 10 in a way that all the firms are identical. As presented in Table 7, PCMs are decreasing as the number of firms is increasing. For this illustrative example, the break-up of a large supplier lowers the prices toward competitive levels. Regulatory bodies can use such analysis for break-up and divestiture decisions in order to ensure "just and reasonable" market prices.

Table 7: PCMs for TOU and single pricing Nash-Cournot models for different number of firms

PCMs / No. of firms	*Off-peak*	*Mid-peak*	*On-peak*	*Single*
2 firms (duopoly)	31.6%	31.9%	31.4%	31.4%
3 firms (after break-up)	23.5%	27.4%	27.6%	16.6%
10 identical firms	7.2%	3.5%	1.7%	4.0%

2.4 *Conclusions and Future Research*

This chapter presents TOU pricing models for electricity markets and proposes a policy tool for regulatory bodies (e.g., to examine the market power exerted by suppliers and to forecast TOU/single prices). The model also allows for reconciliation of the differing time scales of responses of producers and consumers to changing prices. TOU pricing and single pricing schemes in different market structures (from perfect competition to monopoly models) are illustrated with realistic data from the Ontario market. A comparative welfare analysis is conducted for the implementation of TOU prices over single pricing for different market structures. It is concluded that for this illustrative example, overall welfare gains for consumers can be realized when the pricing scheme is changed from single pricing to TOU pricing. Moreover, break-up of the larger firm into two or more parts may increase this gain for the consumers and lower the prices closer to the levels of perfect competition structure.

Another advantage of the models presented in this chapter is that they use the VI framework. VI problems, as compared to MCPs, are easier to create and manage.

The proposed models would be useful for jurisdictions (e.g., Ontario) to assess market power issues by regulatory bodies, to forecast future TOU prices and to examine welfare changes. Another use of the model would be to forecast the TOU prices in a market in which the regulator does not set the prices, but it only defines the different intervals of the day for different prices.

By introducing a linearized DC network, line limits, and ramp limits, a more realistic model can be built and the impact of transmission network (e.g., the effect of location) can be examined in detail, such as the market power issues in load pockets (Bompard et al., 2005; Helman and Hobbs, 2010; also refer to the model in chapter 3 with linearized DC network constraints).

3. Time of Use Pricing Models in Electricity Markets on a Linearized DC Network

A wide range of models for Nash-Cournot game in electricity markets with linearized DC network are proposed in literature, see Hobbs (2001), Day et al. (2002), Metzler et al. (2003), Hobbs and Pang (2007), Hobbs and Helman (2010) and references therein. Hobbs (2001) provided a framework for linear complementarity models of Nash-Cournot game in electricity markets with realistic strategic behavior and physical constraints. He presented two models; one for bilateral market[11] and the other for POOLCO[12] based system. In his static model of bilateral market, Hobbs introduced strategic (Cournot) supplier (e.g. generator) firms, ISO (e.g., "grid owner" or regional transmission operator, RTO) with the responsibility of efficient centralized allocation of transmission services and price-taking spatial arbitrage firm with unlimited capacity (e.g., "perfect arbitragers" that eliminates spatial price discrimination by Cournot firms). It has been shown that this model is equivalent to POOLCO type market model with "no arbitrage" (Hobbs and Helman, 2010).

In the Nash-Cournot setting, each supplier firm assumes that other firms will not alter their sales and that their outputs will not significantly alter transmission prices (Bertrand game for transmission). Also, a common assumption in a Nash-Cournot game is that, all firms are aware of the price-quantity relations (i.e., demand function) in the market.

Following Hobbs' (2001) framework, we would like to introduce TOU pricing in Nash-Cournot game setting in electricity markets on a linearized DC network with line

---

[11] A market in which customers bilaterally contract with individual suppliers and arbitragers to provide energy, and an ISO charges wheeling fees upon these transactions (Metzler et al., 2003).

[12] A centrally administered auction in which generators (consumers) sell (buy) their power to (from) the auctioneer (e.g., POOLCO or ISO) at a locationally dependent price. Auctioneer receives the supply and demand bid to maximize the total welfare (Metzler et al., 2003).

limits. In this setting, it is assumed that there are no power losses during transmission, and congestion is the basis for geographical differentiation in pricing. ISO is the owner of the grid and it operates the transmission system (not only the market operator, and not only ensuring that supply equals demand at every hour, as in chapter 2). ISO charges a congestion based fee (e.g., wheeling fee) for transmitting power from an arbitrary hub node to any node. But these fees are exogenous to its problem (i.e., adopting the Nash-Bertrand assumption that it cannot alter the fees it gets). Also, there are no arbitragers in the model presented here and this allows non-cost based price differences to arise. Hence, suppliers can raise prices where competition is weak or demand is inelastic.

The supplier firms have their decision making process based on a bilateral market model. In this process, generation firms bilaterally contract with consumers to deliver electricity and generation firms pay the cost of transmitting power from the point of generation to the point of consumption. The schedule of injections and withdrawals by generation firms are then provided to the ISO, who collects the transmission fees from these firms for their use of the transmission network. Arbitrage can be introduced at the expense of problem size and complexity, but for our purposes (e.g., computation of equilibrium using modified and approximate Dantzig-Wolfe decomposition algorithms) in chapter 4, it remains in future research directions. For the same reason, extensions to this bilateral model such as the ISO operating a spot market in which generation firms (consumers) can unilaterally sell (buy) power at nodal spot prices is included in future research. Most U.S. ISOs in operating day-ahead markets follow this "hybrid" (bilateral and spot market) model and it is supported by U.S. Federal Energy Regulatory Commission (Hobbs and Pang, 2007). For simplicity, we model the market as entirely bilateral.[13]

---

[13] As Hobbs and Pang (2007, p.115) pointed out "In fact, the bulk of transactions take place on a bilateral basis, but a significant amount flows through the spot market." They also mentioned that the incorporation of a spot market is a "straightforward" extension (Hobbs and Pang, 2007, p.119).

3.1 *Inclusion of Linearized DC Network to TOU Pricing Models*

This section presents a multi-firm, single-period (e.g., one month) equilibrium model in electricity markets with TOU pricing and linearized (lossless) DC network constraints. We introduce more complexity with the transmission structure, so to keep the complexity down and focus on transmission, we suppose only single period (e.g., a month). It is straightforward to include several periods as in chapter 2, but at the cost of being harder to read. The model consists of four parts: the ISO's problem, supply side (e.g., firm *f*'s problem), demand side and the market clearing conditions.

In this setting, the ISO's role is to ensure system balance (e.g., hourly demand and supply balance) and to manage congestion (e.g., routing the power through the transmission system). Operation procedures are similar to the ones in chapter 2 with the following differences:

ISO's problem is to efficiently allocate the transmission service. The power is treated as if it is routed through an arbitrary hub node and congestion (e.g., wheeling) based fees are calculated for transmission service at each node (e.g., bus) of the network. Within the market clearing conditions, the load variation within a demand block is enforced to follow the historical pattern (by using the parameters, δ_{jh}, that are explained in section 3.1.4). The aggregate demand of all firms for each demand block j and each node n in the network is distributed among each hour h according to the historical load-duration curve. Therefore, the ISO's congestion based fees also include a penalty/payment scheme (as in chapter 2 models) for enforcing the historical pattern of the load duration curve. Also, different than the models in chapter 2, ISO as the grid owner (or RTO) collects revenue for providing the transmission service (i.e., maximize the value of the transmission capacity). This revenue is used to recover the fixed costs[14] of the grid owner (i.e., ISO in our models).

[14] In actual markets, a monthly or annual access fee is paid to grid owner (e.g., RTO) to recover their fixed costs (e.g., due to revenue requirements for the holders of the transmission rights) (Helman, 2003). An additional unit charge per MW can be included in the models for this purpose, but it is not modeled here.

Symbols for the ISO's and supply side problems are defined in the following list. Symbols for the demand side are defined in section 3.1.3.

Sets

set of generation facilities: $i=1,\ldots,I$

set of demand blocks: $j=1,\ldots,J$ (alias index k)

set of nodes: $n=1,\ldots,N$ (N_d set of demand nodes; N_g: set of generation nodes)

set of hours in a demand block j: $h=1,\ldots,H_j$ (defined by the market regulator)

set of firms: $f=1,\ldots,F$

set of lines: $l=1,..,L$

Parameters

c_{fni} = operating cost per unit of energy for firm f's facility i at node n (\$/MWh)

κ_{fni} = capacity of firm f's facility i at node n (MW)

δ_{njh} = fraction of total energy demand at node n during block j of a month that occurs during hour h

$PTDF_{ln}$ = power transfer distribution factors[15]

T_{l-}, T_{l+} = lower and upper bounds on real power flows through line (interface) l (MW).

Decision variables

z_{fnijh} = the energy flowing from firm f's facility i to demand block j for hour h at generation node n, $n \in N_g$ (MWh)

d_{fnj} = sales by firm f to demand block j at demand node n, $n \in N_d$ (MWh)

p_{nj} = TOU prices (e.g., uniform block prices) for demand block j at node n (off-peak, mid-peak and on-peak, \$/MWh) (a function of d_{fnj} variables)

y_{njh} = net injections from transmission lines into node n for demand block j at hour h (conceptually, power from hub node to node n) (MW)

---

[15] Power transfer distribution factor for node n on line l ($PTDF_{ln}$) describes the per megawatt (MW) impact (e.g., increase or decrease) in flow resulting from 1 MW of power injection at hub node and 1 MW of withdrawal at node n. Summation of such impacts over all nodes gives the total flow on line l. See appendix B for the derivation of these parameters from the non-linear AC power flow equations.

β_{njh} = a congestion based fee (e.g., wheeling fee) for transmitting power from an arbitrary hub node to node n.

3.1.1 ISO's Problem

In (3.1), the ISO (also the grid owner or RTO) maximizes its profit by allocating transmission capacity efficiently. It chooses y_{njh} by naively assuming that it is a price taker for transmission services (i.e., wheeling fees β^*_{njh} are exogenous in its problem). This is equivalent to a competitive market for transmission rights in which suppliers do not exercise market power (Hobbs, 2001). In this market setting, congestion (wheeling) charges are sufficient to ration the use of the transmission network (i.e., transmission service is allocated efficiently and it maximizes social welfare) (Boucher and Smeers, 2001).

$$
\begin{aligned}
&\max_{y} \sum_{n=1}^{N}\sum_{j=1}^{J}\sum_{h=1}^{H_j} \beta^*_{njh} y_{njh} && && \\
&\textit{subject to} && && [dual] \\
&-\sum_{n=1}^{N} PTDF_{\ln} y_{njh} \le T_{l-} && \forall\, l,j,h && [\gamma_{ljh-}] \\
&\sum_{n=1}^{N} PTDF_{\ln} y_{njh} \le T_{l+} && \forall\, l,j,h && [\gamma_{ljh+}] \\
&\sum_{n=1}^{N} y_{njh} = 0 && \forall\, j,h && [\eta_{jh}]
\end{aligned}
\tag{3.1}
$$

where γ_{ljh-}, γ_{ljh+} are the dual variables of the negative and positive power flow through line l for demand block j at hour h (\$/MW) and η_{jh} is the dual variable for flow balance equation. The model (3.1) differs from that of Hobbs (2001) in the inclusion of hours h and demand blocks j.

Consistent with the linear DC approximation, flows through line l are modeled with PTDFs (see appendix B for derivation of these parameters) which are derived based on Kirchhoff's current law (net flow into a node equals zero) and Kirchhoff's voltage law (net voltage drop around any loop in the network is zero). The net MW flow through line l is $\sum_{n=1}^{N} PTDF_{\ln} y_{njh}$. The last constraint is the flow balance constraint at the hub node. Note that y_{njh} variables are not restricted in sign. A positive (negative) y_{njh} means that there is a net flow into (out of) node n from (to) hub node. It is trivial to note that $y_{njh} = 0$ is always a feasible solution to the ISO's problem, because T_{l-} and T_{l+} are positive scalars (Hobbs, 2001).

3.1.2 Supply Side: Firm f's Problem

In the supply side of the model, formulated in (3.2), firm *f* computes its nodal sales and generation to maximize its profit, π_f, i.e., the total revenues of firm *f* minus the total operating cost of firm *f*'s hourly generation by different technologies of production (e.g., nuclear, hydro, coal, gas/oil, indexed by *i*) to meet its sales in different demand blocks (e.g., off-peak, mid-peak, on-peak, indexed by *j*) minus the ISO's congestion (e.g., wheeling) fees.

$$
\begin{aligned}
&\max_{d,z}\ \pi_f = \sum_{n\in N_d}\sum_{j=1}^{J}\left[p_{nj}(\cdot) - \sum_{h=1}^{H_j}\delta_{njh}\beta^*_{njh}\right]d_{fnj} - \sum_{n\in N_g}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{h=1}^{H_j}\left[c_{fni} - \beta^*_{njh}\right]z_{fnijh} \\
&\mathit{subject\ to} \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad [dual] \\
&\sum_{n\in N_d} d_{fnj} - \sum_{n\in N_g}\sum_{i=1}^{I}\sum_{h=1}^{H_j} z_{fnijh} \le 0 \qquad\qquad \forall\, j \qquad\qquad [\rho_{fj}] \\
&z_{fnijh} \le \kappa_{fni} \qquad\qquad\qquad\qquad \forall\, n,i,j,h \qquad\qquad [\mu_{fnijh}] \\
&z_{fnijh} \ge 0 \qquad\qquad\qquad\qquad\quad \forall\, n,i,j,h
\end{aligned}
\tag{3.2}
$$

The first set of constraints ensures that electricity supply of firm *f* is sufficient to meet its sales to demand block *j* overall nodes; at an optimal solution, these constraints are binding equalities. The second set of constraints contains the capacity constraint for each generation facility owned by firm *f* at each node.

We have examined the Nash-Cournot structure where either all firms or some large firms act à la Cournot. In this structure, firms see their knowledge of the dependence of $p_{nj}(\cdot)$ on total market demand (i.e., in firm *f*'s problem, p_{nj} is a function of $d_{fnj} + \sum_{f\neq g} d^*_{gnj}$, where other firms sales, d^*_{gnj}, are exogenous –denoted by superscript *). It is also presumed that the ISO's wheeling fees, β^*_{njh}, are exogenous to firm *f*'s problem (i.e., all firms are "price taker" for transmission services) and yet endogenous in the overall equilibrium model (i.e., they will become endogenous in the MCP or VI formulation of the whole model).

The $[c_{fni} - \beta^*_{njh}]$ term in the objective function is denoting the per unit cost of the firm *f* to transmit power to the hub node. The $\left[p_{nj}(\cdot) - \sum_{h=1}^{H_j} \delta_{njh}\beta^*_{njh}\right]$ term is denoting the per unit revenue of the firm *f* for conveying energy from hub node to the sales node for demand block *j*.[16] Note that d_{fnj} is a variable for demand block *j* (for several hours), and β^*_{njh} is an hourly wheeling fee. δ_{njh} parameters are the connection between these different time scales (e.g., d_{fnj}: sales to demand blocks *j* at demand node *n*; z_{fnijh}: hourly generation output at generation node *n*). Wheeling fee for sales in demand block *j* at demand node *n* is represented for several hours, i.e., $\sum_{h=1}^{H_j} \delta_{njh}\beta^*_{njh}\, d_{fnj}$.

3.1.3 Demand Side

The demand side is represented by demand equations that use only the prices as independent variables. Different than the demand side in section 2.2.4, the lagged demand term, $\boldsymbol{D}_n^{(0)}$, is a parameter in the single period model. A multi-commodity case where each commodity is the electricity demand in different times of day (e.g., demand blocks: on-peak, mid-peak, off-peak) at each node *n*:

$$\boldsymbol{D}_n = \boldsymbol{A}_n + \boldsymbol{B}_n\boldsymbol{P}_n + \boldsymbol{E}_n\boldsymbol{D}_n^{(0)} \tag{3.3}$$

where

$\boldsymbol{A}_n$ = vector of the factors representing non-price effects at node *n*,

$\boldsymbol{D}_n$ = vector of all demands for electricity at node *n* (i.e., on-peak, mid-peak, off-peak demand at node *n*) where $\boldsymbol{D}_n = [d_{nj}]$ and $d_{nj} = \sum_{f=1}^{F} d_{fnj}$.

$\boldsymbol{D}_n^{(0)}$ = vector of all lagged demands for electricity at node *n*

[16] A firm pays $-\beta^*_{njh}$ to transmit power to the hub node from a generator at node *n* and it pays $\sum_{h=1}^{H_j} \delta_{njh}\beta^*_{n'jh}$ to convey power to sales node *n′* from hub node. Generation is charged straightforwardly at the hourly wheeling fee, but consumption is more complicated because it is measured only in a block of hours, so it is charged at the weighted average of the hourly wheeling fees for the hours within the demand block *j*.

$\boldsymbol{P}_n$ = vector of TOU electricity prices at node n (i.e., on-peak, mid-peak, off-peak prices)
$\boldsymbol{B}_n$ = a square matrix of the price coefficients (i.e., own-price and cross-price) for node n
$\boldsymbol{E}_n$ = a square diagonal matrix of the lag coefficients for node n.

$\boldsymbol{B}_n$ is assumed to be invertible for our analyses (i.e., inverse demand functions are well defined).

3.1.4 Market Clearing Conditions

The total transmission service demanded by generators and consumers from the hub to any node n, demand block j, hour h must equal the transmission service the grid provides between these nodes.

$$\delta_{njh}\sum_{f=1}^{F} d_{fnj} - \sum_{f=1}^{F}\sum_{i=1}^{I} z_{fnijh} = y_{njh} \quad \forall\, n, j, h \quad [dual]\colon [\beta_{njh}] \tag{3.4}$$

where β_{njh} is the dual variable for the market clearing condition (i.e., wheeling fees that clear the markets). Note that total generation by all firms for hour h at node n in demand block j is a fraction (δ_{njh}) of total energy sales by all firms to demand block j at node n. Also note that in the special case of $\delta_{njh} = \delta_{jh}$ at all nodes, we can add up (3.4) over all nodes to derive (together with the last constraint of (3.1)):

$$\delta_{jh}\sum_{n\in N_d}\sum_{f=1}^{F} d_{fnj} - \sum_{n\in N_g}\sum_{f=1}^{F}\sum_{i=1}^{I} z_{fnijh} = 0 \quad \forall\, j, h. \tag{3.5}$$

This is similar to the third set of constraints in (2.2) of the chapter 2, which states that the hourly generation at all generation nodes for all different facilities and all firms should meet the total sales of all firms over all demand nodes for every hour h. With this condition, the ISO imposes the historical shape of the load duration curve within the hours of demand block j over all nodes and β_{njh} can be additionally interpreted as a penalty/payment for deviations from the historical shape of the load duration curve.

Notice that demand variations over hours within the demand block j at each node n is modeled with δ_{njh} parameters (when enough data is available for each node). This pattern of variation in demand at node n, δ_{njh}, is imposed on total sales of all firms, not individually for each firm. A firm may produce more or less than $\delta_{njh} \sum_{n\in N_d} d_{fnj}$ in an hour, but its total production for demand block j must meet its sales in demand block j, $\sum_{n\in N_d} d_{fnj}$.

3.2 MCP and VI Formulations for the TOU Pricing Models on a Linearized DC Network

Firstly, we formulate the perfect competition model as a MCP, by writing out the necessary KKT conditions for the ISO's and firm f's problems along with the demand equation and the market clearing conditions:

MCP: Find $d_{fnj}, z_{fnijh}, p_{nj}, \rho_{fj}, \mu_{fnijh}, y_{njh}, \gamma_{ljh-}, \gamma_{ljh+}, \eta_{jh}, \beta_{njh}$ that satisfy

$$
\begin{array}{lll}
d_{fnj} \geq 0 \perp & -p_{nj} + \sum_{h=1}^{H_j} \delta_{njh}\beta_{njh} + \rho_{fj} \geq 0 & \forall f, n, j \\
z_{fnijh} \geq 0 \perp & c_{fni} - \rho_{fj} + \mu_{fnijh} - \beta_{njh} \geq 0 & \forall f, n, i, j, h \\
\rho_{fj} \geq 0 \perp & \sum_{n\in N_d} d_{fnj} - \sum_{n\in N_g} \sum_{i=1}^{I} \sum_{h=1}^{H_j} z_{fnijh} \leq 0 & \forall f, j \\
\mu_{fnijh} \geq 0 \perp & z_{fnijh} \leq \kappa_{fni} & \forall f, n, i, j, h \\
y_{njh}\ free \perp & -\beta_{njh} + \sum_{l=1}^{L} PTDF_{ln}(\gamma_{ljh+} - \gamma_{ljh-}) + \eta_{jh} = 0 & \forall n, j, h \\
\gamma_{ljh-} \geq 0 \perp & -\sum_{n=1}^{N} PTDF_{ln} y_{njh} \leq T_{l-} & \forall l, j, h \\
\gamma_{ljh+} \geq 0 \perp & \sum_{n=1}^{N} PTDF_{ln} y_{njh} \leq T_{l+} & \forall l, j, h \\
\eta_{jh}\ free \perp & \sum_{n=1}^{N} y_{njh} = 0 & \forall j, h \\
\beta_{njh}\ free \perp & \delta_{njh} \sum_{f=1}^{F} d_{fnj} - \sum_{f=1}^{F} \sum_{i=1}^{I} z_{fnijh} = y_{njh} & \forall n, j, h \\
 & \sum_{f=1}^{F} d_{fnj} = a_{nj} + \sum_{k=1}^{K} b_{njk} p_{nk} + e_{njj} \sum_{f=1}^{F} d_{fnj}^{(0)} & \forall n, j
\end{array}
\tag{3.6}
$$

where a_{nj} and p_{nj} are the j^{th} elements of vectors $\boldsymbol{A}_n$ and $\boldsymbol{P}_n$, respectively. Similarly, b_{njk} and e_{njj} are the elements of matrices $\boldsymbol{B}_n$ and $\boldsymbol{E}_n$, respectively. Note that p_{nj} variables are implicitly defined by the d_{fnj} variables. Instead, an explicit inverse demand function can be used for a more compact formulation without p_{nj} variables, but for ease of readability of the formulation, the p_{nj} variables are used.

We can also formulate (3.6) as a VI problem. The feasible set for the VI problem is defined as follows:

$$K = \left\{ d_{fnj}, z_{fnijh}, p_{nj}, y_{njh} \middle| \begin{array}{ll} \sum_{n \in N_d} d_{fnj} - \sum_{n \in N_g} \sum_{i=1}^{I} \sum_{h=1}^{H_j} z_{fnijh} \leq 0 & \forall f, j \\ z_{fnijh} \leq \kappa_{fni} & \forall f, n, i, j, h \\ z_{fnijh} \geq 0 & \forall f, n, i, j, h \\ -\sum_{n=1}^{N} PTDF_{ln} y_{njh} \leq T_{l-} & \forall l, j, h \\ \sum_{n=1}^{N} PTDF_{ln} y_{njh} \leq T_{l+} & \forall l, j, h \\ \sum_{n=1}^{N} y_{njh} = 0 & \forall j, h \\ \delta_{njh} \sum_{f=1}^{F} d_{fnj} - \sum_{f=1}^{F} \sum_{i=1}^{I} z_{fnijh} = y_{njh} & \forall n, j, h \\ \sum_{f=1}^{F} d_{fnj} = a_{nj} + \sum_{k=1}^{K} b_{njk} p_{nk} + e_{njj} \sum_{f=1}^{F} d_{fnj}^{(0)} & \forall n, j \end{array} \right\}$$

In the feasible set K, the first six constraints are from the firm f's and ISO's problems and the last two equations are the market clearing condition and the linear distributed lagged demand equation, respectively.

The VI problem for the perfect competition model is as in (3.7). Note that for d_{fnj}, z_{fnijh} and y_{njh} the corresponding elements of $\boldsymbol{F}$ are the partial derivatives of the objective functions of firm f's problem (3.2) and the ISO's problem (3.1).

Find $\left(d^*_{fnj}, z^*_{fnijh}, p^*_{nj}, y^*_{njh}\right) \in K$ such that

$$-\sum_{f=1}^{F}\sum_{n\in N_d}\sum_{j=1}^{J}\left(p^*_{nj} - \sum_{h=1}^{H_j}\delta_{njh}\beta_{njh}\right)\left(d_{fnj} - d^*_{fnj}\right) + \sum_{f=1}^{F}\sum_{n\in N_g}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{h=1}^{H_j}\left(c_{fni} - \beta_{njh}\right)\left(z_{fnijh} - z^*_{fnijh}\right) - \sum_{n=1}^{N}\sum_{j=1}^{J}\sum_{h=1}^{H_j}\beta_{njh}\left(y_{njh} - y^*_{njh}\right) \geq 0 \quad \forall \left(d_{fnj}, z_{fnijh}, p_{nj}, y_{njh}\right) \in K \tag{3.7}$$

Also note that the β_{njh} terms are cancelled out in (3.7) (i.e., due to market clearing condition), and a more compact form is derived as follows:

Find $\left(d^*_{fnj}, z^*_{fnijh}, p^*_{nj}\right) \in K$ such that

$$-\sum_{f=1}^{F}\sum_{n\in N_d}\sum_{j=1}^{J} p^*_{nj}\left(d_{fnj} - d^*_{fnj}\right) + \sum_{f=1}^{F}\sum_{n\in N_g}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{h=1}^{H_j} c_{fni}\left(z_{fnijh} - z^*_{fnijh}\right) \geq 0 \quad \forall \left(d_{fnj}, z_{fnijh}, p_{nj}\right) \in K \tag{3.8}$$

The VI problem (3.8) in primal variables has the KKT conditions listed in (3.6) and hence is equivalent to the MCP (3.6)[17].

We only provide the VI formulation for Nash-Cournot market structure, for ease of representation. It is straightforward to derive its equivalent MCP formulations as in the perfect competition case. The Nash-Cournot market structure has the same feasible set K. The VI problem for the Nash-Cournot model is formulated as in (3.9):

[17] There is a minor technicality in the derivation of (3.6) from the KKT conditions of (3.8), and this can be similarly explained as in section 2.3.

Find $(d^*_{fnj}, z^*_{fnijh}, p^*_{nj}) \in K$ such that

$$-\sum_{f=1}^{F}\sum_{n\in N_d}\sum_{j=1}^{J}(p^*_{nj}+\theta^*_{fnj})(d_{fnj}-d^*_{fnj}) + \sum_{f=1}^{F}\sum_{n\in N_g}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{h=1}^{H_j} c_{fni}(z_{fnijh}-z^*_{fnijh}) \geq 0 \quad \forall\,(d_{fnj}, z_{fnijh}, p_{nj}) \in K \tag{3.9}$$

where the term $p^*_{nj}+\theta^*_{fnj}$ is the marginal revenue for firm f at node n and demand block j, and θ^*_{fnj} is the "extra" marginal revenue term. This marginal revenue term is derived from the partial derivative of the objective function in (3.2), with respect to d_{fnj}, when the firm is aware of the price-quantity relation of the distributed lagged demand equation:

$$\frac{\partial \pi_f}{\partial d_{fnj}} = p_{nj} + \frac{\partial p_{nj}}{\partial d_{fnj}} d_{fnj} = p_{nj} + \theta_{fnj} \tag{3.10}$$

where $\left[\frac{\partial p_{nj}}{\partial d_{fnj}}\right] = \boldsymbol{B}_n^{-1}$.

Note that the congestion fees charged by the ISO are neither included in the VI formulation (3.9) nor in the marginal revenue term (3.10), because they are cancelled out in the overall formulation.

3.3 *Illustrative Results for the 66-Bus Ontario Test System*

For our illustrations, we have used the network data for the Ontario electricity system obtained from Wong (2005). It is a scaled down 66-bus version of the full system that is used by the IESO to simulate Ontario's electricity network on a small scale. It is composed of 53 loads, 171 transmission lines, and 12 generators.

This example is a single period (month of July) version of the illustrations in section 2.3.1, e.g., it consists of four types of generation facilities (nuclear, hydro, coal, gas/oil), and three demand blocks (on-peak, mid-peak and off-peak electricity), two firms with firm 1 owning almost 70% of the available generation capacity and firm 2 owning the remainder of the capacity. Each generator unit at a generation node is assigned to a single firm (i.e., firms are spatially competing over the network). The models have been reduced in size by using a representative weekday of the month, as in section 2.3.1.

Because there are no available nodal demand parameters (e.g., a_{nj}, b_{njk}, e_{njj}) for this test system (nor is nodal demand and price data available), we have constructed plausible demand parameters (suitable for computational tests) for each node by the following procedure: We first use the demand parameters from section 2.3.1 to solve a perfect competition TOU pricing model with linearized DC network constraints as in (3.7) by using aggregated demand equations over all nodes[18] without any transmission line limits. We obtain a solution for nodal demands and then we linearize the price and lag elasticities from Celebi and Fuller (2007) around this nodal solution to obtain price ($\boldsymbol{B}_n$) and lag ($\boldsymbol{E}_n$) coefficient matrices for each node. Finally, we compute $\boldsymbol{A}_n$ using the nodal solution with these coefficient matrices. In other words, we have calibrated the demand parameters in the perfect competition model of Celebi and Fuller (2007) to obtain the same results for aggregated (by node) and disaggregated demand equations of the perfect competition model in (3.7). Also we use the special case of $\delta_{njh} = \delta_{jh}$ at all nodes for the fractions that represent the pattern of variation within a demand block j, e.g., identical pattern of variation at all nodes.

$PTDF_{ln}$ for line-node pairs are calculated using the line data from Wong (2005) (i.e., hub node is an arbitrary demand node). We consider a transmission network with congestion, i.e., limits in transmission capacities. The transmission capacities are from Wong (2005), and they are realistically set, but not necessarily correct. Also, we provide results for two types of competition: perfect competition model (3.8) and Nash-Cournot model (3.9).

The models are coded in GAMS using a Windows 2003 Server, Dual Core AMD Opteron (8 CPUs) 2.6 GHz computer with 32GB memory. The VI problems are

[18] We use the inverse demand functions, $\boldsymbol{P}_n = \boldsymbol{B}^{-1}\left(\sum_{n \in N_d} \boldsymbol{D}_n - \boldsymbol{A} - \boldsymbol{E} \sum_{n \in N_d} \boldsymbol{D}_n^0\right)$, where demands are summed over nodes (e.g., aggregated) and $\boldsymbol{D}_n^0 = \boldsymbol{D}_0 \,/\, N$, with the parameters $\boldsymbol{B}^{-1}$, $\boldsymbol{A}$, $\boldsymbol{E}$ and $\boldsymbol{D}^0$ from section 2.3.1.

automatically converted to MCP models by the GAMS\EMP framework and solved by PATH.

In Table 8, we provide the TOU prices at selected nodes for perfect competition model and the Nash-Cournot model. For the perfect competition model, the TOU prices are same for all nodes. This is because of the aforementioned data calibration to obtain the demand parameters for the test system. Therefore, in our illustrations, we focus on the Nash-Cournot model. At nodes (e.g., buses) numbered 2106 and 3107, the TOU prices are relatively smaller compared to other nodes of the Nash-Cournot model. The main reason is that, there is usually congestion in four major lines (e.g., congestion from buses numbered 2100 to 2106, 4105 to 3107, 3107 to 4905 and 8112 to 8104, see Table 11) and the congestion fees (e.g., wheeling fees from hub node to these nodes) are significantly higher than the ones for other nodes in the system for any demand block j and hour h (see Table 12).

Table 8: TOU prices (p_{nj}) for perfect competition and Nash-Cournot market structures at selected nodes ($/MWh)

Bus #	***Off-peak***	***Mid-peak***	***On-peak***
PERFECT COMPETITION			
for all nodes	15.40	27.91	28.60
NASH-COURNOT			
1	32.90	58.53	63.39
101	32.88	58.55	63.41
344	32.90	58.53	63.39
1104	32.87	58.55	63.43
2100	32.91	58.53	63.39
2106	26.51	43.63	48.22
3107	22.80	36.53	37.59
4105	32.90	58.53	63.39
5105	32.90	58.53	63.38
6400	32.90	58.53	63.39
7100	32.91	58.54	63.34
8104	32.91	58.55	63.33
8112	33.26	59.02	62.84
8114	33.20	58.93	62.96
9112	33.26	59.02	62.84
9302	33.26	59.02	62.84
9311	33.26	59.02	62.84

Table 9: PCMs for Nash-Cournot TOU pricing models at selected nodes

Bus #	Off-peak	Mid-peak	On-peak
1	53.2%	52.3%	54.9%
101	53.2%	52.3%	54.9%
344	53.2%	52.3%	54.9%
1104	53.1%	52.3%	54.9%
2100	53.2%	52.3%	54.9%
2106	41.9%	36.0%	40.7%
3107	32.5%	23.6%	23.9%
4105	53.2%	52.3%	54.9%
5105	53.2%	52.3%	54.9%
6400	53.2%	52.3%	54.9%
7100	53.2%	52.3%	54.8%
8104	53.2%	52.3%	54.8%
8112	53.7%	52.7%	54.5%
8114	53.6%	52.6%	54.6%
9112	53.7%	52.7%	54.5%
9302	53.7%	52.7%	54.5%
9311	53.7%	52.7%	54.5%
AVERAGE	**52.7%**	**51.5%**	**54.0%**

In Table 9, we present the PCMs at selected nodes, defined as the difference between Nash-Cournot price and perfect competition price divided by the market Nash-Cournot price. Also, PCMs averaged over all nodes are presented. Note that the Nash-Cournot structure with network constraints has higher PCMs than the ones presented in chapter 2 (see Table 2).

The welfare analyses (i.e., change in consumers', producers' and total surpluses for selected nodes) for changing from Nash-Cournot model to perfect competition model (i.e., reference case) is summarized in Table 10 for selected nodes. It is noted that at all nodes, total surpluses and consumers' surpluses increase whereas suppliers' surpluses decreases. Also, note that the lag effect on welfare (as explained in Appendix A) is not included in the analyses, because the results are for a single-period model (e.g., no lag demand). The consumers' surplus gains as percentage of the total energy bill at selected nodes are presented in the last column of Table 10. Similar to chapter 2 results, changing from Nash-Cournot model to perfect competition model under TOU pricing scheme has some significant effect on these gains in our illustrations.

Table 10: Change in consumers', suppliers' and total surpluses over selected nodes due to change from Nash-Cournot oligopoly to perfect competition and consumers' surpluses as percent of total energy bill (last column)

Change in / Bus #	Consumers' Surplus (CS)	Suppliers' Surplus	Total Surplus	CS as % Total Energy Bill
1	$258,130	-$150,744	$107,385	63.1%
101	$181,040	-$105,724	$75,316	63.1%
344	$249,657	-$145,796	$103,860	63.1%
1104	$159,690	-$93,257	$66,433	63.1%
2100	$266,248	-$155,485	$110,763	63.1%
2106	$297,924	-$200,042	$97,882	42.7%
3107	$39,605	-$28,275	$11,330	27.3%
4105	$264,086	-$154,223	$109,864	63.1%
5105	$263,518	-$153,890	$109,627	63.1%
6400	$263,698	-$153,996	$109,701	63.1%
7100	$261,730	-$152,841	$108,888	63.1%
8104	$261,529	-$152,723	$108,807	63.1%
8112	$66,887	-$38,861	$28,026	63.5%
8114	$74,578	-$43,351	$31,227	63.5%
9112	$66,887	-$38,861	$28,026	63.5%
9302	$66,887	-$38,861	$28,026	63.5%
9311	$66,887	-$38,861	$28,026	63.5%
OVERALL	**12,274,294**	**-8,342,742**	**3,931,552**	**62.1%**

Table 11: Dual variables ($\gamma_{ljh-/+}$) of the negative and positive power flow through line l for demand block j at hour h ($/MW) (legend: ~~**generation nodes**~~, **off-peak**, *mid-peak*, *on-peak*)

Hours / Bus n to m	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	*11*	*12*
2100 to 2106	9.98	9.98	9.98	9.98	9.98	9.98	9.98	22.10	22.10	22.10	23.58	23.58
4105 to 3107	19.65	19.65	19.65	19.65	19.65	19.65		38.35	38.35	35.61	40.62	40.62
3107 to ~~4905~~							19.65			2.74		
8112 to 8104											1.53	1.53
Hours / Bus n to m	*13*	*14*	*15*	*16*	*17*	*18*	*19*	*20*	*21*	*22*	**23**	**24**
2100 to 2106	23.58	23.58	23.58	23.58	23.58	22.10	22.10	22.10	22.10	22.10	3.90	9.98
4105 to 3107	40.62	40.62	40.62	40.62	40.62		38.35	38.35	38.35	38.35		19.65
3107 to ~~4905~~						38.35					19.65	
8112 to 8104	1.53	1.53	1.53	1.53	1.53							

In Table 11, the duals associated with the line limits are represented. As aforementioned, some lines are congested, which in turn has an effect on congestion fees (i.e., power is wheeled through the hub node to node n, hence any congestion on the route has an impact on this fee). Congestion (e.g., wheeling) fees (β_{njh}) are shown in Table 12.

Table 12: Congestion (e.g., wheeling) fees (β_{njh}) for transmitting power from an arbitrary hub node to node n (\$/MWh) (legend: ~~**generation nodes**~~, **off-peak**, <u>*mid-peak*</u>, *on-peak*)

Bus # \ Hours	**1**	**2**	**3**	**4**	**5**	**6**	**7**	<u>*8*</u>	<u>*9*</u>	<u>*10*</u>	*11*	*12*
2106	-9.98	-9.98	-9.98	-9.98	-9.98	-9.98	-9.98	-22.10	-22.10	-22.10	-23.58	-23.58
~~**2962**~~	-9.98	-9.98	-9.98	-9.98	-9.98	-9.98	-9.98	-22.10	-22.10	-22.10	-23.58	-23.58
3107	-19.37	-19.37	-19.37	-19.37	-19.37	-19.37	0.27	-37.82	-37.82	-35.08	-40.06	-40.06
~~**4905**~~	-12.73	-12.73	-12.73	-12.73	-12.73	-12.73	-12.73	-24.85	-24.85	-24.85	-26.33	-26.33
8112											1.21	1.21
8114											1.13	1.13
~~**9103**~~											1.21	1.21
9112											1.21	1.21
9302											1.21	1.21
9311											1.21	1.21

Bus # \ Hours	*13*	*14*	*15*	*16*	*17*	<u>*18*</u>	<u>*19*</u>	<u>*20*</u>	<u>*21*</u>	<u>*22*</u>	**23**	**24**
2106	-23.58	-23.58	-23.58	-23.58	-23.58	-22.10	-22.10	-22.10	-22.10	-22.10	-3.90	-9.98
~~**2962**~~	-23.58	-23.58	-23.58	-23.58	-23.58	-22.10	-22.10	-22.10	-22.10	-22.10	-3.90	-9.98
3107	-40.06	-40.06	-40.06	-40.06	-40.06	0.53	-37.82	-37.82	-37.82	-37.82	0.27	-19.37
~~**4905**~~	-26.33	-26.33	-26.33	-26.33	-26.33	-24.85	-24.85	-24.85	-24.85	-24.85	-12.73	-12.73
8112	1.21	1.21	1.21	1.21	1.21							
8114	1.13	1.13	1.13	1.13	1.13							
~~**9103**~~	1.21	1.21	1.21	1.21	1.21							
9112	1.21	1.21	1.21	1.21	1.21							
9302	1.21	1.21	1.21	1.21	1.21							
9311	1.21	1.21	1.21	1.21	1.21							

Table 13: Weighted sum of congestion fees ($\sum_{h=1}^{H_j} \delta_{njh}\beta_{njh}$) and consumers' payments received by firms net of weighted sum of congestion fees ($p_{nj} - \sum_{h=1}^{H_j} \delta_{njh}\beta_{njh}$)

	$\sum_{h=1}^{H_j} \delta_{njh}\beta_{njh}$			$p_{nj} - \sum_{h=1}^{H_j} \delta_{njh}\beta_{njh}$		
Bus #	***Off-peak***	***Mid-peak***	***On-peak***	***Off-peak***	***Mid-peak***	***On-peak***
2106	-9.24	-22.10	-23.58	35.75	65.74	71.80
3107	-14.62	-32.48	-40.06	37.43	69.02	77.65
8112			1.21			61.63
8114			1.13			61.83
9112			1.21			61.63
9302			1.21			61.63
9311			1.21			61.63

We have computed the weighted sum of congestion fees that is deducted from the consumers' payments to firms in Table 13. For nodes numbered 2106 and 3107, firm 2's facilities that are located at nodes numbered 2962 (nuclear) and 4905 (hydro) receive incentives, because they reduce congestion and follow the historical shape of the load duration curve for their overall sales (i.e., node 2962 is only connected to node 2106 and node 4905 is only connected to nodes 3107 and 4105). On the other hand, firm 2's facility at node numbered 9103 is penalized by ISO, as it causes congestion and does not follow the historical pattern of the load-duration curve (i.e., node 9103 is connected to nodes 9112, 9302 and 9311; and through node 9112, it is connected to nodes 8112, which is also connected to 8114).

3.4 Conclusions and Future Research

There are many ways in which this basic framework can be extended. Arbitragers that erases non-cost based price differences among nodes can be introduced in the model (Hobbs and Helman, 2004; Hobbs and Pang, 2007). In addition to providing transmission service, ISO can perform the arbitrage activity by introducing unrestricted arbitrage variables in the objective function of its problem with the objective function coefficients equal to $p_{nj}^* - \sum_{h=1}^{H_j} \delta_{njh}\beta_{njh}$ (Hobbs et al., 2008). Alternatively, ISO can reallocate power to maximize consumers' surplus (Yao et al., 2005). Another possibility is to add a separate price-taking arbitrage firm that maximizes its revenues (e.g., arbitrage amounts multiplied by the price difference in different nodes) and pays congestion fee (e.g., β_{njh}) subject to arbitrage balance constraints (e.g., total injections and withdrawals of arbitraged energy sum to zero). The alternative extensions to bilateral model can be used to simulate prices for a POOLCO based system.

We have experimented with MCP formulation of a realistic 66-bus network for Ontario market (a single month perfect competition model with only 2-firms and line limits) and it took more than 5 hours to reach equilibrium (there were around 16,500 variables and conditions). In such a case, decomposition methods (e.g., DW or Benders decomposition for VI problems) may surmount difficulties that may arise in computation of equilibrium. We explore DW decomposition in chapter 4.

4. Dantzig-Wolfe Decomposition Method for VI Problems with Applications to Two Models in Energy Markets

This chapter is devoted to the DW decomposition for VI problems and its application to two models in energy markets. First, we present a general overview of decomposition methods from the literature. Then, we introduce the DW decomposition algorithms, namely exact, modified and approximate DW algorithms. Numerical investigations are performed on two models in energy markets. One of these models is a single-period (month) TOU pricing electricity market equilibrium model with linearized DC network constraints from chapter 3. We relate the general VI forms of the DW subproblem and master problem to this problem after the introduction of each algorithm in section 4.2. The other model is a realistic two-region energy equilibrium model for Canada from Fuller and Chung (2005).

4.1 Decomposition Methods for Variational Inequality Problems

Fuller and Chung (2005, p.304) address the main challenges to adapt optimization based decomposition to VI problems as:

1) Definition of the forms of subproblems and master problem, and communication among them;
2) Definition of a valid stopping (convergence) condition;
3) Proof of convergence of the algorithm.

There are various decomposition algorithms for optimization problems. DW and Benders decomposition are the most well known of those algorithms and they have been extended to VI problems. In addition to DW and Benders decomposition methods, other algorithms such as Lagrangean methods, simplicial decomposition, cobweb decomposition and partitionable decomposition are presented in the next subsections. This section is a general overview of these methods and algorithms. Section 4.2 is devoted to DW decomposition algorithm and its modifications, as well as applications to TOU pricing models of chapter 3.

4.1.1 Lagrangean Methods

Lagrangean methods are based on relaxing the constraints that prevent the separability of the problem into subproblems (e.g., subproblems that are easy to solve) by using dual multipliers that penalize violations of these constraints. The method then iterates by updating these dual multipliers by subgradient methods, cutting plane methods or bundle methods (Conejo et al., 2006). Lagrangean methods for VI problems include the work by He et al. (1999, 2004), Han and Sun (2003) and Auslender and Teboulle (2000).

4.1.2 Dantzig-Wolfe Decomposition

DW decomposition is a special case of the column generation principle (Dantzig and Wolfe, 1961). The idea behind column generation is to algorithmically generate profitable variables (columns or proposals) to improve the objective function. An example of this technique is for the cutting stock problem where a column generation method is based on the solution of knapsack problems. In the dual formulation, column generation corresponds to cutting plane methods (i.e., adding a constraint).

DW decomposition takes advantage of a block angular structure of the constraint set to decompose the problem into a master problem and one or more simpler (e.g., easy to solve) subproblems. The master problem is defined over the convex combination of the subproblem columns (proposals). The subproblem(s) employs a pricing out mechanism of the simplex method for column generation.

Chung et al. (2003) employs the DW decomposition method for a class of economic equilibrium models. They decompose the asymmetric economic equilibrium problem into one equilibrium master problem (but it was not stated as a VI) and one LP subproblem. They distinguish between supply and demand variables as a special case of the model represented in Ahn and Hogan (1982) for the purpose of decomposition. Proposals from the subproblem to the master problem only include supply variables and the master problem has its own demand variables in addition to the weight variables for the convex combinations of the proposals. They provide a convergence criterion and conclude that the convergence occurs in a finite number of iterations, since the subproblem is a LP model with finite extreme points and there is no repetition of proposals (i.e., finite number of proposals from subproblem).

Fuller and Chung (2005) introduce the DW decomposition for VI problems and show that their algorithm is convergent under usefully general conditions. The algorithm begins with relaxing a VI by removing complicating (linking) constraints and solving the subproblem. The master problem, a restricted version of the original VI problem (i.e., restricted to the convex combinations of the proposals from the subproblem), is then solved and computed dual solutions (prices) are passed to the subproblem. The algorithm proceeds between master and subproblem until a scalar quantity called the "*convergence gap*" is close enough to zero. A small negative value of the convergence gap indicates that the current master problem solution is very close to equilibrium. They illustrate this convergence behavior by a realistic two-region energy equilibrium model for Canada.

An extension of this work has been developed for multiregional and multicommodity economic equilibrium models. In the former work, DW decomposition is applied to decompose the multiregional model by region and a two-region energy equilibrium example is given to illustrate the convergence behavior and some useful properties of the algorithm (Chung et al., 2006). The latter one combines the Jacobi idea of PIES algorithm with DW decomposition of VI problems (Chung and Fuller, 2005). In this case, each equilibrium subproblem contains variables corresponding to not only demand for its own commodity, but also demands for other commodities. Since there is no information about the supply of other commodities in the subproblem, an equilibrium solution cannot be found. Using the Jacobi idea of PIES algorithm (by fixing the other commodities demand at the most recent values computed by the master problem), an equilibrium solution is found for the single commodity subproblem (which is now integrable and can be converted to a nonlinear optimization problem). The convergence results are provided under general useful conditions. Also, they do not need to assume that own-price influence on demand is more important than cross-price effects as is required for the convergence proof of the PIES method (see Ahn and Hogan, 1982).

Chung and Fuller (2010) further extended their study on DW decomposition of VI problems by approximating the subproblem in a variety of ways. These approximations allow them to decompose a multicommodity economic equilibrium model into separate subproblems for each commodity. They have illustrated these approximation techniques on the same model of Chung et al. (2006) and concluded that these approximations can reduce the number of iterations required to achieve convergence. Moreover, they have pointed out that decomposition by commodity can be useful for model management purposes. They have also provided (in an online appendix to the paper) that the convergence theorems and proofs hold for these approximations of the subproblem.

4.1.3 Benders Decomposition

In Benders decomposition, the complicating variables are fixed at the most recent master problem solution and the resulting subproblem is solved iteratively (Geoffrion, 1972). Based on the subproblem's associated dual, a cutting plane (i.e., a linear inequality) that "cuts off' the current solution point is found by the algorithm. This cut is added to the constraint set of the master problem, which is then re-solved.

Lawphongpanich and Hearn (1990) apply Benders Decomposition principle for VI problems. They define the subproblem VI by fixing some of the variables (complicating variables) that are determined by the master problem VI. However, they do not provide any proof of convergence of the algorithm for general VI problems. Also, the master problem of their proposed method requires a point-to-set mapping which requires strong assumptions on the mapping for convergence of the algorithm. Chung et al. (2003) states that the non-convexity of the transformed feasible region makes this problem very hard to solve. Lawphongpanich and Hearn (1990) implement their algorithm for a partially asymmetric traffic assignment problem where they use an approximation of the master problem, a LP, by using a cutting plane approach that adds a new constraint at each iteration with the dual information from the most recent subproblem VI. Their numerical example is to demonstrate that the "algorithm is implementable and has potential as an effective method for decomposing large VI problems" (Lawphongpanich and Hearn, 1990, p.245).

Fuller and Chung (2008) apply Benders decomposition and provide convergence results and proofs for a useful class of VI problems. Their algorithm is mainly based on DW decomposition of VI problems by Fuller and Chung (2005) and Chung et al. (2006). They apply a DW decomposition procedure to a dual of the given VI. By converting the dual forms of DW master and subproblems to their primal forms, they derive the Benders master and subproblems. At each iteration, information from the latest subproblem adds a new cut to the Benders master problem and a scalar convergence gap parameter is calculated. They prove that the algorithm makes progress when the convergence gap is negative and under mild conditions it approaches zero in the limit of many iterations. They also show that if it has reached a non-negative value, a solution is obtained (under more restrictive conditions, i.e., strict monotonicity).

Gabriel and Fuller (2010) presented a Benders decomposition method for solving a general two-stage stochastic complementarity problem (or VI), which is an extension of Fuller and Chung (2008) study. They have applied the Benders decomposition method (where decomposition corresponds to scenarios in their stochastic complementarity problem) on Hobbs' (2001) electricity market equilibrium model, but for which stochastic elements are included. Their numerical investigations have shown substantial improvements in computation time over the extended form of the original model. They have also extended the theoretical results for DW and Benders decomposition methods of Fuller and Chung (2005; 2008) to the case where the master problem has its own variables (e.g., scenario independent variables can be retained in the master problem).

4.1.4 Simplicial Decomposition

Simplicial decomposition is also a special case of the column generation principle, where the column generating subproblem is formed by approximating (often by a linear function) the objective function of the original problem (Patriksson, 1999). It has been applied to the VI problems, mostly in the context of traffic assignment problems. But this method does not allow for decomposition, since the subproblem is still subject to the whole feasible set. Hence, the block angular feasible set cannot be decomposed. Also, the subproblem defined by this method is a LP, not a VI.

4.1.5 Cobweb Decomposition

The cobweb algorithm is an iterative method that passes price or quantity approximations between subproblems (e.g., demand and supply models) and the algorithm stops when the price and quantity approximations do not change. But this decomposition algorithm may diverge even in one-dimensional examples (Ahn and Hogan, 1982).

Murphy and Mudrageda (1998) study the convergence of a cobweb algorithm motivated by the experience with the National Energy Modeling System[19] (NEMS) for the U.S. Department of Energy. They use step functions to approximate the supply and demand curves and employ a weighted average scheme[20] for the convergence of the cobweb algorithm. They infer from their NEMS experience that their algorithm converges even when they have crude implementation and rough approximations.

4.1.6 Partitionable Decomposition

Partitionable decomposition takes advantage of the separable structure of the feasible region of some VI problems. Such problems, referred to as partitionable VI problems, are a class of VI problems in which the constraints define a Cartesian product

[19] For evaluation of energy modeling in U.S. Department of Energy see Gabriel et al. (2001), Murphy and Mudrageda (1998) and Murphy and Shaw (1995).

[20] Weighted average of the last two iterates is used instead of the most recent iterate.

of feasible sets. This type of VI problem can be decomposed into subproblems (coupled VI problems of small dimensions) and these subproblems can be solved by linear or nonlinear approximations (Nagurney, 1993; Patriksson, 1994). However, the main assumption in this decomposition method is the separable structure of the feasible set. This method is not applicable for block angular feasible sets.

4.2 *Dantzig-Wolfe Decomposition Method for TOU Pricing Models with Transmission Network Constraints*

In this section, algorithms for the DW decomposition and modifications to this algorithm including an approximation of the master problem solution are explained in detail. Applications to TOU pricing models of chapter 3 are also provided. For simplicity of representation, a single period (month) TOU pricing model will be given as an example throughout the illustrations, but a multi-period model extension is straightforward. Decomposition algorithms are applied to this example. Firstly, we introduce the general problem setting as well as the subproblem and master problem definitions.

We summarize the main results of Fuller and Chung (2005), using a slightly different notation and following their presentation closely. We consider a VI problem with a feasible set defined by two sets of constraints. We distinguish one of these constraint sets as complicating constraints, e.g., when they are relaxed a VI subproblem is formed (and it may or may not be decomposable, but it is easier to solve). Convex combinations of solutions of the subproblem with the complicating constraints form a master problem. We first define the feasible set for the original VI as follows. All vectors are considered to be column vectors. The feasible set is $K = \{x \in R^n | g(x) \geq 0, h(x) \geq 0\}$, where g is a mapping from R^n to R^m such that g_i is concave and continuously differentiable for all $i = 1, \ldots, m$, and h is a mapping from R^n to R^l such that h_i is concave and continuously differentiable for all $i = 1, \ldots, l$. The constraints $h(x) \geq 0$ represent the complicating constraints. The vector function G maps R^n to R^n. The original VI is defined as follows:

$$VI(K,G)\text{: find } x^* \in K \text{ such that } G^T(x^*)(x - x^*) \geq 0 \quad \forall\, x \in K. \tag{4.1}$$

The feasible set for the subproblem is defined by relaxing the complicating constraints in K and it is represented as: $\bar{K} = \{x \in R^n | g(x) \geq 0\}$. The subproblem at iteration *k+1* is defined with ω^k (the dual variable vector corresponding to the complicating constraints from the previous master problem solved at iteration k) and $\nabla h(x_M^k)$ (the matrix of gradients of h evaluated at the master problem solution x_M^k). The subproblem VI is defined as follows:

$$\begin{aligned} &Sub\text{-}VI^{k+1}(\bar{K}, G - \nabla h^T(x_M^k)\omega^k)\text{: find } x_S^{k+1} \in \bar{K} \text{ such that} \\ &\left(G(x_S^{k+1}) - \nabla h^T(x_M^k)\omega^k\right)^T(x - x_S^{k+1}) \geq 0 \quad \forall\, x \in \bar{K}. \end{aligned} \tag{4.2}$$

The feasible set for the master problem at iteration k is restricted to all convex combinations of the k solutions (or "proposals") that have been calculated by the first k solutions of the subproblem. The complicating constraints need to be satisfied in this set. We use the notation $X^k = [x_S^1, x_S^2, \ldots, x_S^k]$ to represent matrices whose columns are the k proposals collected from the subproblem at each iteration. The weights on the proposals in the convex combination are contained in the vector $\lambda \in R^k$. The feasible set for the master problem is defined as: $\Lambda^k = \{\lambda \in R^k | h(X^k\lambda) \geq 0,\ e^{kT}\lambda = 1,\ \lambda \geq 0\}$, where $e^k \in R^k$ is a vector whose k entries are all one. Since the convex combination $X^k\lambda$ of solutions from the subproblem enforces the constraint set $g(x) \geq 0$, there is no need to explicitly mention this set of the constraints in the feasible set of the master problem. For brevity, we sometimes use the notation x_M^k to denote the solution of the master problem, λ^k, in terms of the original x variables: $x_M^k = X^k\lambda$. Note that $x_M^k \in K \subseteq \bar{K}$ (i.e., the feasible region of the subproblem contains the original problem's feasible region, since it is a "relaxed" version of the original problem without complicating constraints). Finally, the master problem at iteration k is defined as:

$$Master\text{-}VI^k(\Lambda^k, H^k)\text{: find } \lambda^k \in \Lambda^k \text{such that } H^{kT}(\lambda^k)(\lambda - \lambda^k) \geq 0 \quad \forall\, \lambda \in \Lambda^k \tag{4.3}$$

where the mapping H^k from R^k to R^k is used to denote the condition $G^T(x^*)(x - x^*) \geq 0$ in terms of λ, i.e., $H^{kT}(\lambda^k) = G^T(X^k\lambda)X^k$.

As an alternative notation, the feasible set for the master problem is also denoted by K^k:

$$K^k = \{x \in R^k | h(x) \geq 0, x \in conv(X^k)\}$$

where $conv(X^k)$[21] stands for the convex hull of the points represented by X^k. The relationship among the feasible sets can be summarized as:

$$K^1 \subseteq K^2 \subseteq \cdots \subseteq K^k \subseteq K^{k+1} \subseteq \cdots \subseteq \bar{K} \subseteq K.$$

The master problem is more compactly defined as:

$$\textit{Master-VI}^k(K^k, G)\text{: find } x^* \in K \text{ such that } G^T(x^*)(x - x^*) \geq 0 \quad \forall\, x \in K^k. \tag{4.4}$$

The DW algorithm uses the following information exchange between master problem and subproblem. The subproblem for $k = 1$, i.e., *Sub-VI*^{1}, is solved with a starting guess of the value of the mapping adjustment, such as $\nabla h^T(x_M^0)\omega^0 = 0$, to obtain the proposals to be transferred to the matrix X^k of the master problem, thus enlarging the set Λ^k (or K^k). Then *Master-VI*^{1} is solved to estimate a new dual vector ω^1. Later iterations begin with a subproblem and end with a master problem. After each subproblem is solved, a scalar quantity called the convergence gap is calculated from the solutions of subproblem *Sub-VI*$^{k+1}$ and master problem *Master-VI*^{k}:

$$CG^k = \left(G(x_M^k) - \nabla h^T(x_M^k)\omega^k\right)^T (x_S^{k+1} - x_M^k). \tag{4.5}$$

The algorithm terminates when a predetermined convergence tolerance, $\varepsilon > 0$, is reached, e.g., $|CG^k| < \varepsilon$.

Fuller and Chung (2005) have provided useful convergence and existence results under the following assumptions:

1) $\bar{K}$ is bounded.
2) Each component of $h(x)$ and $g(x)$ is concave and continuously differentiable.
3) G is continuous.
4) Subproblem and master problem are feasible at each iteration.
5) G is strictly monotone.[22]

[21] $conv(X^k) = \{\lambda \in R^k | x = X^k\lambda,\ e^{kT}\lambda = 1,\ \lambda \geq 0\}$

[22] Fuller and Chung (2005) has also introduced an assumption that is applicable to a part of the mapping G, i.e., $G = \begin{pmatrix} -p(d) \\ \nabla c(z) \end{pmatrix}$ where $x = \begin{pmatrix} d \\ z \end{pmatrix}$, $-p(d)$ is strictly monotone and $c(z)$ is a convex function.

Under these assumptions, Fuller and Chung (2005) prove several results, including: if $CG^k < 0$, then x_M^k solves *VI(K,G)*; before convergence, $CG^k < 0$, and $\lim_{k\to\infty} CG^k = 0$. Furthermore, if in Assumption 5, "strictly monotone" is replaced by "strongly monotone," then $\lim_{k\to\infty}\|x_S^{k+1} - x_M^k\| = 0$ (if G is strongly monotone). We have provided the theorems of Fuller and Chung (2005) in Appendix C and the proofs can be found in Fuller and Chung (2005) or Chung and Fuller (2010).

It is worth mentioning that sets of solutions to subproblems are assumed to be bounded as well, since unboundedness of subproblems creates complications by passing unbounded rays to master problem. This can be avoided by imposing upper\lower bounds on x variables in the subproblems that do not already have any upper\lower bounds. Also it is assumed that the master problem has an equilibrium solution at every iteration. Infeasibility of the master problem can be avoided by introducing artificial variables in the complicating constraints with high cost coefficients, i.e., $h_i(x) + a_i \geq 0 \quad \forall\, i$ with $a_i \geq 0$ (Chung et al., 2006). However, determining the value of these high cost coefficients (e.g., "big-M" values) for artificial variables may cause some problems. If they are too small, positive artificial variables may be observed in the solution, and if they are too large, numerical problems due to poor scaling may arise. In practice, modeler's insight is important in determining the "big-M" values, i.e., they are the bounds on the dual variables of the complicating constraints (Fuller and Chung, 2010).

In the next subsections, we present the algorithms for the DW decomposition method in detail.

4.2.1 Exact Dantzig-Wolfe Decomposition Algorithm

In this subsection, we provide the exact DW algorithm and relate it to the VI formulation of TOU pricing models in chapter 3. The VI problem for the perfect competition model of TOU pricing is as in (3.7). To relate (3.7) to the general VI form (4.1), subproblem (4.2) and master problem (4.4), we have

$$x = \begin{bmatrix} d_{fnj} \\ z_{fnijh} \\ y_{njh} \\ p_{nj} \end{bmatrix} \forall f, n, i, j, h;\; G(x) = \begin{bmatrix} -p_{nj} \\ c_{fni} \\ 0 \\ 0 \end{bmatrix} \forall f, n, i;\; \omega^k = \left[\beta_{njh}^k\right] \forall n, j, h;$$

$$-\nabla h^T\left(x_M^k\right)\omega^k = \begin{bmatrix} \sum_{h=1}^{H_j} \delta_{njh}\beta_{njh}^k \\ -\beta_{njh}^k \\ -\beta_{njh}^k \\ 0 \end{bmatrix} \forall n, j, h; \quad X^k = \begin{bmatrix} D_{fnj}^k \\ Z_{fnijh}^k \\ Y_{njh}^k \\ P_{nj}^k \end{bmatrix} = \begin{bmatrix} d_{fnj}^1, d_{fnj}^2, \dots, d_{fnj}^k \\ z_{fnijh}^1, z_{fnijh}^2, \dots, z_{fnijh}^k \\ y_{njh}^1, y_{njh}^2, \dots, y_{njh}^k \\ p_{nj}^1, p_{nj}^2, \dots, p_{nj}^k \end{bmatrix} \forall f, n, i, j, h.$$

The vector x contains the variables $d_{fnj}, z_{fnijh}, y_{njh}$ and p_{nj} for all f, n, i, j and h, and the elements of the vector-valued mapping $G(x)$ are as follows: $-p_{nj}$ is the element of G that corresponds to d_{fnj}; c_{fni} is the element of G that corresponds to z_{fnijh}; and the elements of G that corresponds to y_{njh} and p_{nj} are zero. Note that for d_{fnj} and z_{fnijh}, the corresponding elements of G are the partial derivatives of the objective function of firm f's problem (3.2), but the terms involving the dual variables β_{njh}^k are left out. Because, these dual terms β_{njh}^k are cancelled out in the original VI formulation and therefore, they do not appear in the mapping G. The feasible set K is as in (3.8) and the feasible set for the subproblem and the master problem are as follows:

$$\bar{K} = \left\{ d_{fnj}, z_{fnijh}, p_{nj}, y_{njh} \middle| \begin{array}{ll} \sum_{n\in N_d} d_{fnj} - \sum_{n\in N_g}\sum_{i=1}^{I}\sum_{h=1}^{H_j} z_{fnijh} \le 0 & \forall f, j \\ z_{fnijh} \le \kappa_{fni} & \forall f, n, i, j, h \\ z_{fnijh} \ge 0 & \forall f, n, i, j, h \\ -\sum_{n=1}^{N} PTDF_{ln} y_{njh} \le T_{l-} & \forall l, j, h \\ \sum_{n=1}^{N} PTDF_{ln} y_{njh} \le T_{l+} & \forall l, j, h \\ \sum_{n=1}^{N} y_{njh} = 0 & \forall j, h \\ \sum_{f=1}^{F} d_{fnj} = a_{nj} + \sum_{k=1}^{K} b_{njk} p_{nk} + e_{njj} \sum_{f=1}^{F} d_{fnj}^{(0)} & \forall n, j \end{array} \right\}$$

$$K^k = \left\{ d_{fnj}, z_{fnijh}, p_{nj}, y_{njh} \in conv(X^k) \middle| \delta_{njh} \sum_{f=1}^{F} d_{fnj} - \sum_{f=1}^{F} \sum_{i=1}^{I} z_{fnijh} = y_{njh} \; \forall \; n, j, h \right\}$$

Note that the complicating constraints corresponding to $h(x)$ are the market clearing conditions (3.4) and they are the constraints in the feasible set of the master problem, K^k. Once they are eliminated the ISO's problem and firms' problems can be separated (i.e., since p_{nj} depends on other firm's sales, firms' problems cannot be decomposed, hence, it is cast as a single equilibrium subproblem). In our computational illustrations, the subproblem was left as one large subproblem and it was not split into separate subproblems for ISO and firms; because, in almost all of our experiments, most computation time is spent for the master problem, we chose to focus on the modification/approximation of the master problem.

The exact decomposition algorithm is stated below. Note that a "null matrix" is a matrix with no columns.

Exact DW Algorithm

Step 0: Set *k=0*. Choose β_{njh}^0 and $\varepsilon > 0$. Set $\left[D_{fnj}^0, Z_{fnijh}^0, Y_{njh}^0, P_{nj}^0\right]^T$ to the null matrix and $CG^k = -\infty$

Step 1: Increment $k \longleftarrow k+1$. Solve *Sub-VI^k* with $\beta_{njh}^k = \beta_{njh}^{k-1}$ and place the solution

$$x_S^k = \left[d_{fnj}^k, z_{fnijh}^k, y_{njh}^k, p_{nj}^k\right]^T in \begin{bmatrix} D_{fnj}^k \\ Z_{fnijh}^k \\ Y_{njh}^k \\ P_{nj}^k \end{bmatrix} = \begin{bmatrix} D_{fnj}^{k-1}, d_{fnj}^k \\ Z_{fnijh}^{k-1}, z_{fnijh}^k \\ Y_{njh}^{k-1}, y_{njh}^k \\ P_{nj}^{k-1}, p_{nj}^k \end{bmatrix}.$$

If *k=1* then go to Step 2; else

If $|CG^{k-1}| < \varepsilon$, then STOP; else

go to Step 2.

Step 2: Solve *Master-VI^k*. Record β_{njh}^k, the dual of the market clearing condition; record $\left[d_{fnj}^k, z_{fnijh}^k, y_{njh}^k, p_{nj}^k\right]^T$ for calculation of CG^k). Go to Step 1.

The exact DW algorithm is implemented using VI formulation for the subproblem and the master problem in GAMS\EMP framework and solved by the

PATH solver. However, even without any line limits, this decomposition method does not converge after 48 hours of computation time. We have observed that, for the ISO's variables, all the added columns were used at least once during the algorithm. This suggests the idea to include the ISO's problem in the master problem instead of the subproblem.

4.2.2 Modified Algorithm: ISO's Problem in Master Problem

In this subsection, we present a modified DW decomposition algorithm and we show that the subproblem in this algorithm is very similar to the TOU pricing models in chapter 2 (e.g., models without network constraints).

Since the market clearing conditions (3.4) are eliminated from the subproblem's feasible set $\bar{K}$, in the exact DW algorithm, the ISO's problem can be separated from the subproblem and it can alternatively be moved into the master problem. Therefore, in this modified DW algorithm, the y_{njh} variables only appear in the master problem (i.e., the master problem has its own variables as in Gabriel and Fuller (2010) and they have shown that all the convergence results for DW decomposition holds for this case).

For the exact DW algorithm, we observe in our experiments that the proposals from subproblems have almost never satisfied the $\sum_{n=1}^{N} y_{njh} = 0$ constraints in the master problem and therefore, the artificial variables usually remain positive in the master problem for many iterations.

To enforce feasibility for the master problem and to produce better proposals from subproblems, we add extra constraints to the subproblem's feasible set $\bar{K}$. These extra constraints are obtained by summing the complicating constraints over all nodes, e.g., $\delta_{jh} \sum_{n \in N_d} \sum_{f=1}^{F} d_{fnj} - \sum_{n \in N_g} \sum_{f=1}^{F} \sum_{i=1}^{I} z_{fnijh} = \sum_{n=1}^{N} y_{njh}$. The right hand side of this equation, $\sum_{n=1}^{N} y_{njh}$, is equal to zero in the master problem, hence we derive the extra constraints by setting the right hand side to zero:

$$\delta_{jh} \sum_{n \in N_d} \sum_{f=1}^{F} d_{fnj} - \sum_{n \in N_g} \sum_{f=1}^{F} \sum_{i=1}^{I} z_{fnijh} = 0 \quad \forall\, j, h. \tag{4.6}$$

In our experiments, these extra constraints do not complicate the subproblem (e.g., subproblems are very easy to solve within a second of CPU time). They also provide better proposals for the master problem (e.g., by satisfying the $\sum_{n=1}^{N} y_{njh} = 0$ constraints).

Also note that these extra constraints added to the feasible set $\overline{K}$ are very similar to the 3rd set of constraints in (2.1) (e.g., the ISO's problem in chapter 2, without network constraints). In fact, the subproblem for the modified DW algorithm is almost same as the models of chapter 2 (see section 2.3, p. 28). The only differences are the representation of network nodes in the subproblem (e.g., nodal generation, nodal sales and nodal prices) and the adjustment in the mapping, $G(x) - \nabla h^T(x_M^k)\omega^k$.

To relate the modified algorithm's subproblem to the general form of subproblem in (4.2), we have:

$$x = \begin{bmatrix} d_{fnj} \\ z_{fnijh} \\ p_{nj} \end{bmatrix} \ \forall f, n, i, j, h; \quad G(x) = \begin{bmatrix} -p_{nj} \\ c_{fni} \\ 0 \end{bmatrix} \ \forall f, n, i; \quad \omega^k = \left[\beta_{njh}^k\right] \ \forall n, j, h;$$

$$-\nabla h^T(x_M^k)\omega^k = \begin{bmatrix} \sum_{h=1}^{H_j} \delta_{njh}\beta_{njh}^k \\ -\beta_{njh}^k \\ 0 \end{bmatrix} \forall n, j, h.$$

The feasible set for the subproblem is as follows:

$$\overline{K} = \left\{ d_{fnj}, z_{fnijh}, p_{nj} \left| \begin{array}{ll} \sum_{n\in N_d} d_{fnj} - \sum_{n\in N_g}\sum_{i=1}^{I}\sum_{h=1}^{H_j} z_{fnijh} \le 0 & \forall f, j \\ z_{fnijh} \le \kappa_{fni} & \forall f, n, i, j, h \\ z_{fnijh} \ge 0 & \forall f, n, i, j, h \\ \delta_{jh} \sum_{n\in N_d}\sum_{f=1}^{F} d_{fnj} - \sum_{n\in N_g}\sum_{f=1}^{F}\sum_{i=1}^{I} z_{fnijh} = 0 & \forall j, h \\ \sum_{f=1}^{F} d_{fnj} = a_{nj} + \sum_{k=1}^{K} b_{njk}p_{nk} + e_{njj}\sum_{f=1}^{F} d_{fnj}^{(0)} & \forall n, j \end{array} \right. \right\}$$

In this subproblem, the vector x contains the variables d_{fnj}, z_{fnijh} and p_{nj} for all f, n, i, j and h, and the elements of the vector-valued mapping $G(x)$ are as follows: $-p_{nj}$

is the element of G that corresponds to d_{fnj}; c_{fni} is the element of G that corresponds to z_{fnijh}; and the element of G that corresponds to p_{nj} is zero. Note that the feasible set for the subproblem does not include any constraints of the ISO's problem. These constraints are included in the master problem. Also the subproblem does not include any y_{njh} variables in the VI formulation (i.e., the subproblem has no variables for the linearized DC network constraints).

For the modified algorithm, we relate the modified algorithm's master problem to the master problem in (4.4) as follows:

$$x = \begin{bmatrix} d_{fnj} \\ z_{fnijh} \\ y_{njh} \\ p_{nj} \end{bmatrix} \ \forall f,n,i,j,h; \quad G(x) = \begin{bmatrix} -p_{nj} \\ c_{fni} \\ 0 \\ 0 \end{bmatrix} \ \forall f,n,i;$$

$$X^k = \begin{bmatrix} D_{fnj}^k \\ Z_{fnijh}^k \\ P_{nj}^k \end{bmatrix} = \begin{bmatrix} d_{fnj}^1, d_{fnj}^2, \dots, d_{fnj}^k \\ z_{fnijh}^1, z_{fnijh}^2, \dots, z_{fnijh}^k \\ p_{nj}^1, p_{nj}^2, \dots, p_{nj}^k \end{bmatrix} \forall f,n,i,j,h.$$

The feasible set for the master problem is as follows:

$$K^k = \left\{ d_{fnj}, z_{fnijh}, p_{nj} \in conv(X^k), y_{njh} \middle| \begin{array}{ll} -\sum_{n=1}^{N} PTDF_{ln} y_{njh} \le T_{l-} & \forall l,j,h \\ \sum_{n=1}^{N} PTDF_{ln} y_{njh} \le T_{l+} & \forall l,j,h \\ \sum_{n=1}^{N} y_{njh} = 0 & \forall j,h \\ \delta_{njh} \sum_{f=1}^{F} d_{fnj} - \sum_{f=1}^{F} \sum_{i=1}^{I} z_{fnijh} = y_{njh} & \forall n,j,h \end{array} \right\}$$

For the master problem, the vector x contains the variables $d_{fnj}, z_{fnijh}, y_{njh}$ and p_{nj} for all f,n,i,j and h, and the elements of the vector-valued mapping $G(x)$ are as follows: $-p_{nj}$ is the element of G that corresponds to d_{fnj}; c_{fni} is the element of G that corresponds to z_{fnijh}; and the element of G that corresponds to p_{nj} and y_{njh} are zero. Note that, in the feasible set of the master problem $d_{fnj}, z_{fnijh}, p_{nj} \in conv(X^k)$ and y_{njh}

are the variables of the master problem only (e.g., subproblem only provides proposals for d_{fnj}, z_{fnijh}, p_{nj} variables).

The modified DW algorithm (when the ISO's problem is in the master problem) is stated below.

Modified DW Algorithm

Step 0: Set *k=0*. Choose β^0_{njh} and $\varepsilon > 0$. Set $\left[D^0_{fnj}, Z^0_{fnijh}, P^0_{nj}\right]^T$ to the null matrix and $CG^k = -\infty$

Step 1: Increment $k \leftarrow k+1$. Solve *Sub-VI^k^* with $\beta^k_{njh} = \beta^{k-1}_{njh}$ and place the solution $x^k_S = \left[d^k_{fnj}, z^k_{fnijh}, p^k_{nj}\right]^T in \begin{bmatrix} D^k_{fnj} \\ Z^k_{fnijh} \\ P^k_{nj} \end{bmatrix} = \begin{bmatrix} D^{k-1}_{fnj}, d^k_{fnj} \\ Z^{k-1}_{fnijh}, z^k_{fnijh} \\ P^{k-1}_{nj}, p^k_{nj} \end{bmatrix}$.

If *k=1* then go to Step 2; else

If $|CG^{k-1}| < \varepsilon$, then STOP; else

go to Step 2.

Step 2: Solve *Master-VI^k^*. Record β^k_{njh}, the dual of the market clearing condition; record $\left[d^k_{fnj}, z^k_{fnijh}, p^k_{nj}\right]^T$ for calculation of CG^k). Go to Step 1.

With this algorithm and without any line limits, the decomposition method converges in only one iteration. However, with line limits, the modified algorithm takes around four and a half hours to converge and more than 99% of computation time is spent on the master problem. This is an improvement over the exact DW algorithm (which cannot converge to equilibrium solution in a reasonable time). However, we seek a further improvement in computation times by an approximate solution of the master problem rather than an exact solution of it, as explained in the next subsection.

4.2.3 Approximate Algorithm for the Solution of the Master Problem

In this subsection we will provide an approximation method for the solution of the master problem in the DW decomposition algorithm for the TOU pricing models of chapter 3. This algorithm is essentially the same as the modified DW algorithm. We also provide the convergence results and proofs for this approximation method. First, we define the approximation method and its properties in a general setting.

An approximation to the mapping $G(x)$ in the master problem is considered. We approximate the master problem mapping by $\tilde{G}(x;\varphi^k)$, which relies on parameters in the vector φ^k which has the same dimensions as x (e.g., these parameters can include the most recent solution of the master problem, x_M^{k-1}). To gain some computational advantage $\tilde{G}$ is chosen so that the approximate master problem VI is easy to solve (e.g., in section 4.2.4, it becomes a nonlinear programming problem –NLP). It is required that the approximate $\tilde{G}$ should satisfy two properties. The first property replaces Assumption 5 of the section 4.2.

Approximation Properties:

1) One of the following is true:
   (a) $\tilde{G}(x;\varphi^k)$ is strictly monotone in x;
   (b) $\tilde{G}(x;\varphi^k)$ is the gradient of a convex function.
2) If $\varphi^k = x_M^k$, then $\tilde{G}(x_M^k;\varphi^k) = G(x_M^k)$.

The second property is required for the convergence verification only. Different than Chung and Fuller (2010), the original mapping $G(x)$ is *not* preserved in the master problem. This approximation in the mapping of the master problem also requires redefinition of CG^k, because it is evaluated by the mapping $\tilde{G}(x;\varphi^k)$ of the master problem. The subproblem is the same as in (4.2). The master problem with the approximate mapping is as follows:

Master-VI^k(K^k, $\tilde{G}(x;\varphi^k)$): find $x^* \in K^k$ such that $\tilde{G}^T(x^*;\varphi^k)(x - x^*) \geq 0 \quad \forall\, x \in K^k$. (4.6)

The algorithm begins by solving the subproblem for $k=1$, and by choosing an initial guess for φ^k; then *Master- VI^1* is solved. Later iterations begin with a subproblem and end with a master problem. The approximate convergence gap, $\widetilde{CG}^k$, is calculated by using the approximate mapping $\tilde{G}$:

$$\widetilde{CG}^k = \left(\tilde{G}\left(x_M^k;\varphi^k\right) - \nabla h^T\left(x_M^k\right)\omega^k\right)^T\left(x_S^{k+1} - x_M^k\right). \tag{4.7}$$

Under Assumptions 1 to 4 and the two approximation properties, all of the theorems and proofs in Fuller and Chung (2005) hold for this approximation (see Appendix C for statements of all theorems from Fuller and Chung (2005)), with the exception of theorem 6. We provide a new theorem 6 and its proof. The rest of the proofs for the approximation of the mapping G in the master problem by $\tilde{G}$ holds (e.g., they hold by replacing the mapping G with $\tilde{G}$).

New Theorem 6: If $\widetilde{CG}^k \geq \left(\tilde{G}\left(x_M^k;\varphi^k\right) - G\left(x_M^k\right)\right)^T \left(x_S^{k+1} - x_M^k\right)$ then x_M^k solves *VI(K,G)*, under the strict monotonicity of the mapping $G(x)$.

Proof: We shall show that if x_M^k does not solve *VI(K,G)*, then

$$\widetilde{CG}^k < \left(\tilde{G}\left(x_M^k;\varphi^k\right) - G\left(x_M^k\right)\right)^T \left(x_S^{k+1} - x_M^k\right).$$

By theorem 4 of Fuller and Chung (2005), suppose x_M^k does not solve *VI(K,G)*. Then x_M^k does not solve *Sub-VI*$^{k+1}$, so $x_M^k \neq x_S^{k+1}$. Strict monotonicity of G implies that $\left(G\left(x_M^k\right) - G\left(x_S^{k+1}\right)\right)^T \left(x_M^k - x_S^{k+1}\right) > 0$, since $x_M^k \neq x_S^{k+1}$.

Also x_S^{k+1} solves *Sub-VI*$^{k+1}$*(*$\bar{K}, G - \nabla h^T\left(x_M^k\right)\omega^k$*)* and $x_M^k \in \bar{K}$, it follows that

$$\left(G\left(x_S^{k+1}\right) - \nabla h^T\left(x_M^k\right)\omega^k\right)^T\left(x_M^k - x_S^{k+1}\right) \geq 0.$$

Adding this last inequality to the strict inequality, we get:

$$\left(G\left(x_M^k\right) - \nabla h^T\left(x_M^k\right)\omega^k\right)^T\left(x_M^k - x_S^{k+1}\right) > 0$$

Now adding the term $\left(\tilde{G}\left(x_M^k;\varphi^k\right) - G\left(x_M^k\right)\right)^T \left(x_M^k - x_S^{k+1}\right)$ to both sides of the inequality and multiplying by -1 yield:

$$\left(\tilde{G}\left(x_M^k;\varphi^k\right) - \nabla h^T\left(x_M^k\right)\omega^k\right)^T\left(x_S^{k+1} - x_M^k\right) < \left(\tilde{G}\left(x_M^k;\varphi^k\right) - G\left(x_M^k\right)\right)^T \left(x_S^{k+1} - x_M^k\right),$$

i.e., $\widetilde{CG}^k < \left(\tilde{G}\left(x_M^k;\varphi^k\right) - G\left(x_M^k\right)\right)^T \left(x_S^{k+1} - x_M^k\right)$. ■

We know that the difference term, $\tilde{G}\left(x_M^k;\varphi^k\right) - G\left(x_M^k\right)$, in new theorem 6 would vanish as $\varphi^k \to x_M^k$, because of approximation property 2. In the numerical results of section 4.2.4, $\tilde{G}$ is chosen to be a symmetric approximation of G, using $\varphi^k = x_M^{k-1}$ in the resulting NLP that approximates the VI master problem. It is well known that such a sequential NLP algorithm has $x_M^{k-1} \to x_M^k$ under weak conditions without decomposition (see Nagurney (1993), p.40). We check that with just one NLP per master problem iteration, this convergence takes place, but if not, then it would be possible to take extra NLP steps in a master problem iteration. In either case, $\varphi^k \to x_M^k$ as k increases.

This approximation in the decomposition algorithm allows more flexibility than the original DW algorithm for VI problems. One example of an advantage of this flexibility is that, as the subproblem of DW decomposition passes new proposals (e.g., columns) to the master problem, the master problem size grows and it may become computationally challenging to find a solution. Using an approximation method for the solution of the master problem may overcome this challenge, e.g., instead of solving the master problem VI with an asymmetric mapping G, an integrable mapping $\tilde{G}(x_M^k;\varphi^k)$ can be used to form a convex optimization problem, where commercial NLP solvers are faster and robust than current VI or MCP solvers. This may provide computational gains over the original DW decomposition algorithm.

Now we can state the approximate DW decomposition algorithm for the TOU pricing models of chapter 3.

Approximate DW Algorithm

Step 0: Set k=0. Choose β_{njh}^0 and φ^0. Set $\left[D_{fnj}^0, Z_{fnijh}^0, P_{nj}^0\right]^T$ to the null matrix and $\widetilde{CG}^k = -\infty$

Step 1: Increment $k \leftarrow k+1$. Solve *Sub-VI*^{k} with $\beta_{njh}^k = \beta_{njh}^{k-1}$ and place the solution

$$x_S^k = \left[d_{fnj}^k, z_{fnijh}^k, p_{nj}^k\right]^T in \begin{bmatrix} D_{fnj}^k \\ Z_{fnijh}^k \\ P_{nj}^k \end{bmatrix} = \begin{bmatrix} D_{fnj}^{k-1}, d_{fnj}^k \\ Z_{fnijh}^{k-1}, z_{fnijh}^k \\ P_{nj}^{k-1}, p_{nj}^k \end{bmatrix}.$$

If k=1 then go to Step 2; else

If $\widetilde{CG}^{k-1} \geq \left(\tilde{G}(x_M^{k-1};\varphi^{k-1}) - G(x_M^{k-1})\right)^T (x_S^k - x_M^{k-1})$, then STOP; else go to Step 2.

Step 2: Solve *Master-VI*[k] with approximate mapping $\tilde{G}(x;\varphi^k)$ with $\varphi^k = x_M^{k-1}$.[23] Record β_{njh}^k, the dual of the market clearing condition; record $\left[d_{fnj}^k, z_{fnijh}^k, p_{nj}^k\right]^T$ (for calculation of $\widetilde{CG}^k$). Go to Step 1.

As stated above, the approximate DW algorithm does not need any convergence tolerance $\varepsilon > 0$ for the stopping condition, $\widetilde{CG}^{k-1} \geq \left(\tilde{G}\left(x_M^{k-1};\varphi^{k-1}\right) - G\left(x_M^{k-1}\right)\right)^T \left(x_S^k - x_M^{k-1}\right)$. However, in our experiments we have found that this stopping condition can be too strict (although as $k \to \infty$, this condition is satisfied in theory).

Note that, since $\bar{K}$ is bounded, it follows that $\left\|x_S^k - x_M^{k-1}\right\| \leq D$, where D is the maximum distance between points in $\bar{K}$. Therefore, the convergence requirement is bounded above:

$$\left(\tilde{G}\left(x_M^{k-1};\varphi^{k-1}\right) - G\left(x_M^{k-1}\right)\right)^T \left(x_S^k - x_M^{k-1}\right) \leq \left\|\tilde{G}\left(x_M^{k-1};\varphi^{k-1}\right) - G\left(x_M^{k-1}\right)\right\| D.$$

We require the factor, $\left\|\tilde{G}\left(x_M^{k-1};\varphi^{k-1}\right) - G\left(x_M^{k-1}\right)\right\|$, to be as small as possible, because the judgement about stopping can be made essentially as in the exact DW algorithm: if $\widetilde{CG}^{k-1} < 0$, then continue (by theorem 5 of Fuller and Chung (2005), see Appendix C) and if $\widetilde{CG}^{k-1} \geq 0$, then stop. From the approximation property 2, if $\varphi^{k-1} = x_M^{k-1}$, then the factor above is zero. Therefore, we have used a two-part stopping condition in our experiments. In step 0 of the algorithm, we choose tolerances $\varepsilon > 0$ and $\xi > 0$. In step 1 of the algorithm, if $\widetilde{CG}^{k-1} \geq -\varepsilon$ and $max\left|x_M^{k-1} - \varphi^{k-1}\right| < \xi$, then the algorithm is required to stop. We have also verified that this stopping condition enforces the factor, $\left\|\tilde{G}\left(x_M^{k-1};\varphi^{k-1}\right) - G\left(x_M^{k-1}\right)\right\|$, to be close to zero (see Figure 1 in section 4.2.4).

4.2.4 Numerical Results

We illustrate the algorithms explained in section 4.2.2 and 4.2.3 for the TOU pricing models with linearized DC network constraints, as detailed and illustrated in

---

[23] There are other possibilities for choosing φ^k, e.g., weighted average of previous m master problem solutions.

chapter 3. The models are coded in GAMS using a Windows 2003 Server, Dual Core AMD Opteron (8 CPUs) 2.6 GHz computer with 32GB memory. The VI problems are solved by the GAMS\EMP framework and the PATH solver, and the convex optimization problems are solved by the NLP solver CONOPT3. The MCP models are also solved by the PATH solver in GAMS.

We first solve the original models (perfect competition and Nash-Cournot cases) using the MCP and VI formulations. Computational results are summarized in the following tables (CPU times are PATH solver time only and excludes model generation times).

Table 14: Computational Results for the Original Perfect Competition Model with Different Formulations

	No line limits		***With line limits***	
Formulation	*MCP*	*VI*	*MCP*	*VI*
CPU time (sec)	48.0	55.4	18,240.3	69.9
No. of iterations	489	488	163,876	735
No. of Equ./Var.	16,500	16,662	16,500	16662

Table 15: Computational Results for the Original Nash-Cournot Model with Different Formulations

	No line limits		***With line limits***	
Formulation	*MCP*	*VI*	*MCP*	*VI*
CPU time (sec)	47.5	57.5	22,672.0	54.5
No. of iterations	814	429	256,281	483
No. of Equ./Var.	16,824	17,310	16,824	17,310

For the no line limits case, the differences in the computational results for different formulations are very small. But with the line limits, the VI formulation has a substantial advantage. The MCP formulations take around 5 to 7 hours to reach the equilibrium solution, where as the VI models are solved around a minute of CPU time. GAMS/EMP framework, in fact, converts the VI formulation into an equivalent MCP formulation and while doing that some additional equation/variable pairs are added (note the slight increase in number of equation/variable pairs). This is most probably the reason for a faster computation of the equilibrium. Since there is no technical documentation about the GAMS/EMP framework yet (other than a general guide in (Ferris et al., 2009), at this point, we can presume that these added equation/variable pairs make the VI formulation stronger than its pure MCP counterpart. As pointed out in Garcia et al. (2003), this increase in dimension improves the formulation (e.g., for integer programming and for certain structured LPs and NLPs, see Garcia et al. (2003) for details and references). Hence, we have used the VI formulation and GAMS/EMP framework for models in the rest of our computational results.

In our computations, the subproblems are retained as a single equilibrium problem in all algorithms. This most likely added very little to the total solution time, since solving the subproblem with different β_{njh}^{k}values for the TOU pricing models is very fast (less than a second at each DW iteration). In place of computing CG^{k} or $\widetilde{CG}^{k}$ at each iteration of the DW algorithms, we have calculated a relative convergence gap, RCG^{k} or $\widetilde{RCG}^{k}$, for each iteration as a percent of the equilibrium value of producer's surplus, $\sum_{f,n,j} p_{nj}(d^{*})d_{fnj}^{*} - \sum_{f,n,i,j,h} c_{fni} z_{fnijh}^{*}$. Hence, $RCG^{k}\left(\widetilde{RCG}^{k}\right)$ gives an economically meaningful idea about the magnitude of $CG^{k}\,(\widetilde{CG}^{k})$ relative to the equilibrium value of producer's surplus.

For all algorithms, we have set $\varepsilon = 0.01\%$, $\xi = 0.1\%$, $\varphi^{0} = x^{*}$ (where x^{*} is the equilibrium solution for the perfect competition model without any line limits) and $\beta_{njh}^{0} = 0$ (but for Nash-Cournot model β_{njh}^{0} is set to equilibrium solution, β_{njh}^{*}, of perfect competition model solution with line limits). In the master problem, artificial variables are added to the constraint set with large cost coefficients (e.g., big-M). After some experimentation, big-M values are set to 325. Also in the subproblem, large upper\lower bounds are set for all x variables that do not already have any upper\lower bounds.

Within the DW algorithms, computations for subproblem and master problem start from their equilibrium solution found in the previous iteration. This may be computationally advantageous for the iterations that use slightly modified data from the previous iteration (e.g., it may reduce the number of iterations required for convergence) (Murphy et al., 1988).

We were not able to reach an equilibrium solution for the exact DW algorithm for the TOU pricing models with 66-bus network (with or without line limits) within a reasonable time framework (e.g., we terminate the algorithm after 48 hours of CPU time). This is expected, because once the difficult constraints are relaxed, the link between the ISO's problem and the firms' problems is also disconnected. Hence, proposals from the subproblem do not satisfy the relaxed constraints and artificial variables become positive in most of the iterations (i.e., no feasible solution is found). For smaller test systems (e.g., 3-bus system), we have experienced this problem, however, the exact DW algorithm has found a feasible solution after 30 iterations and it has converged to the equilibrium solution after 84 iterations for a 3-bus example. It should be noted that unless there are very few iterations, the exact DW algorithm may work poorly, perhaps because there are many complicating (linking) constraints (e.g., in our illustrations there are $24 \times 66 = 1584$ of these constraints). Hence, other strategies to reduce the number of complicating constraints are also important (e.g., calculating excess supply/demand from subproblems and determining a new price, β_{njh}^{k}, at each iteration without solving the master problem).

The modified DW algorithm, without line limits case, converges in only one iteration. This is not surprising, because without line limits, the congestion based wheeling fees would be zero and using a starting guess of $\beta_{njh}^{0} = 0$ would provide the equilibrium solution for the subproblem, e.g., $\left[d_{fnj}^{1}, z_{fnijh}^{1}, p_{nj}^{1}\right]^{T} = [d_{fnj}^{*}, z_{fnijh}^{*}, p_{nj}^{*}]^{T}$. Also, because of the extra constraints in the feasible set of the subproblem, this solution satisfies $\sum_{n=1}^{N} y_{njh} = 0$ constraints in the master problem.

However, with line limits, the modified algorithm takes 176 iterations to converge (with four and a half hours of CPU time). Therefore, we seek better computational results with the approximation of the master problem in DW decomposition algorithm.

We illustrate the approximate DW algorithm for the TOU pricing models with linearized DC network constraints as in chapter 3. In this model, instead of using the asymmetric inverse demand function, $p_{nj}(d)$ (where $d = [d_{fnj}]$), we define a symmetric inverse demand function, $\tilde{p}_{nj}(d; \varphi^k)$, where φ^k equals the d variables from the most recent solution of the master problem, d_M^{k-1}.

$$\tilde{p}_{nj}(d; \varphi^k) = p_{nj}(\varphi^k) + \tilde{B}^{-1}(d - \varphi^k) = p_{nj}(d) + (\tilde{B}^{-1} - B^{-1})(d - \varphi^k),$$

where $\tilde{B} = \frac{1}{2}(B + B^T)$.

The approximate DW algorithm can also be used to apply one step of PIES algorithm to the master problem by modifying it slightly, i.e., $\tilde{B} = Diag(B)$. We call this approximation, approximate-PIES. This way, we can define a diagonal inverse demand function, $\tilde{p}_{nj}(d; \varphi^k)$, (e.g., with cross-demand variables fixed at the φ^k).

These approximate inverse demand functions in the mapping $\tilde{G}(x; \varphi^k)$ satisfy the approximation properties. Also, $\tilde{G}(x; \varphi^k)$ becomes a gradient of a convex function, e.g., it can be integrated to form a convex optimization problem. To relate (3.8) to the approximate master problem (4.6), we have

$$x = \begin{bmatrix} d_{fnj} \\ z_{fnijh} \\ y_{njh} \\ p_{nj} \end{bmatrix} \ \forall f, n, i, j, h; \quad G(x) = \begin{bmatrix} -p_{nj} \\ c_{fni} \\ 0 \\ 0 \end{bmatrix} \ and \ \tilde{G}(x; \varphi^k) = \begin{bmatrix} -\tilde{p}_{nj}(d; \varphi^k) \\ c_{fni} \\ 0 \\ 0 \end{bmatrix} \ \forall f, n, i;$$

$$X^k = \begin{bmatrix} D_{fnj}^k \\ Z_{fnijh}^k \\ P_{nj}^k \end{bmatrix} = \begin{bmatrix} d_{fnj}^1, d_{fnj}^2, \dots, d_{fnj}^k \\ z_{fnijh}^1, z_{fnijh}^2, \dots, z_{fnijh}^k \\ p_{nj}^1, p_{nj}^2, \dots, p_{nj}^k \end{bmatrix} \forall f, n, i, j, h.$$

In the approximate DW algorithm, we have the two-part stopping condition $\widetilde{CG}^{k-1} \geq -\varepsilon = -0.01\%$ and $max|x_M^{k-1} - \varphi^{k-1}| < \xi = 0.1\%$. We note that the latter condition has made the $\|\tilde{G}(x_M^{k-1}; \varphi^{k-1}) - G(x_M^{k-1})\|$ measure negligible as DW iterations carry on, as depicted in the figure below.

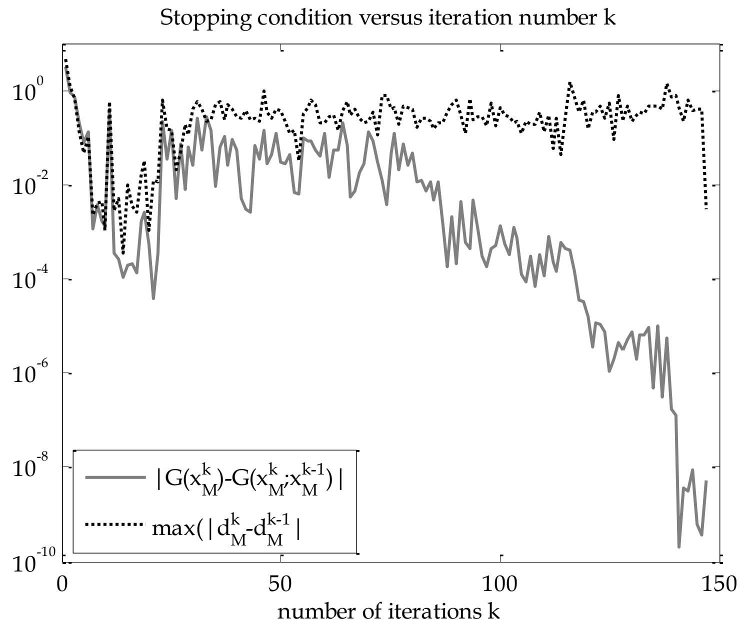

Figure 1: Stopping conditions for the approximate DW algorithm

Without line limits case, the approximate DW algorithm converges, again, in only one iteration. With line limits, it takes around one (two) hours to converge to the equilibrium solution for the perfect competition (Nash-Cournot) model, a huge improvement compared to leaving extra constraints out of $\overline{K}$ and improvement over the modified algorithm with line limits by a factor of four (three).

Computational results for all algorithms are summarized in the following tables.

Table 16: Computational Results for the DW Algorithms for Perfect Competition TOU Pricing Models as in (3.8) (with line limits)

DW Algorithm	***CPU time (sec)***			***DW iterations***
	Sub	*Master*	*Total*	
Exact	n/a	n/a	>48 hours	>300
Modified	122.9	15,971.6	16,094.5	176
Approximate	58.0	3,486.7	3,544.7	147
Approximate (PIES)	56.6	4,170.6	4,227.2	156

Table 17: Computational Results for the DW Algorithms for Nash-Cournot TOU Pricing Models as in (3.9) (with line limits)

DW Algorithm	***CPU time (sec)***			***DW iterations***
	Sub	*Master*	*Total*	
Exact	n/a	n/a	>48 hours	>300
Modified	123.2	22,140.1	22,263.3	230
Approximate	104.1	7,550.4	7,654.5	235
Approximate (PIES)	104.0	7,794.7	7898.7	240

Figure 2 and Figure 3 show the relative convergence gap and the convergence gap values, respectively, at each iteration of the approximate DW algorithms for the perfect competition model (3.8).

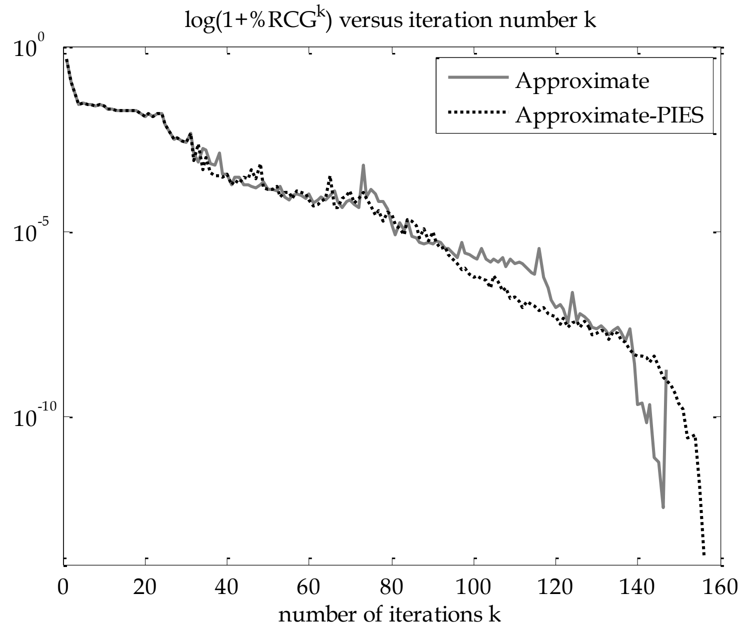

Figure 2: Convergence gap as percent of the producer's surplus at each DW iteration for the approximate DW algorithms

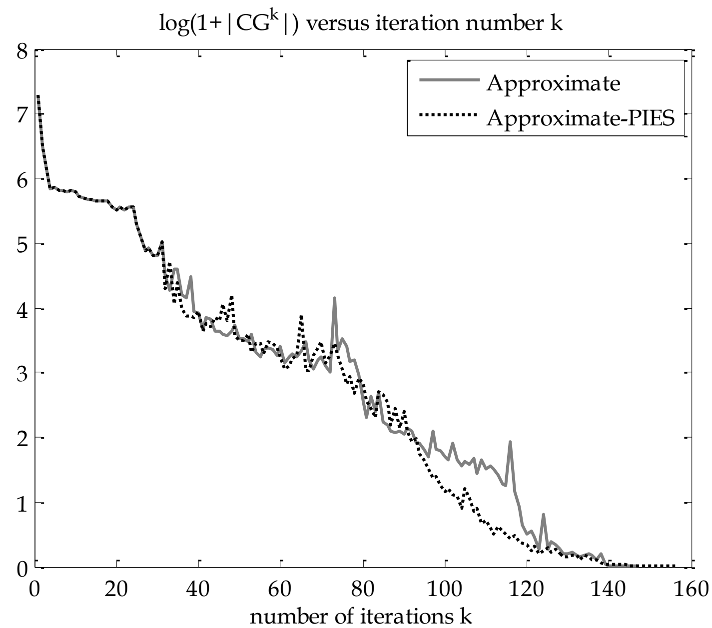

Figure 3: Progress of iterations for the approximate DW algorithms

We can see that the convergence gap is not guaranteed to decrease monotonically as Chung et al. (2006) point out. We have presented the further details (tables for progress of DW iterations) on these algorithms in Appendix D.

Another important finding in the detailed solution is that, artificial variables become all zero after some iterations (e.g., around 20th or 30th iteration). Therefore, they can be dropped from the master problem for the subsequent iterations when a feasible solution is found (e.g., after the 20th or 30th iteration). However, we did experience negligible computational advantage[24] by dropping them after they have reached zero in the DW algorithm.

On the other hand, we have found some computational improvements when they are never included in the master problem of the modified DW algorithm. If the constraint for the sum of the weights for the convex combinations of proposals in the master problem is modified as $e^{kT}\lambda \leq 1$, zero solution becomes feasible. Once feasibility is satisfied[25], artificial variables are not be needed. We have compared the computational results with and without artificial variables for the perfect competition model in the following table.

---

[24] There may be some computational advantages in modeling time for GAMS (e.g., time elapsed to create the model due to less number of variables) but execution times (e.g., CPU times elapsed for PATH solver) are not affected. Because, in our experiments, the solution of the previous iteration is retained for the next iteration (e.g., warm starting). Once all artificial variables become zero for iteration k, the PATH algorithm starts from the zero solution for these variables in iteration *k+1*. Since, zero solution for iteration k is feasible, it is also feasible for iteration *k+1* (due to theorem 5, strict inclusion, of Appendix C).

[25] In the modified DW algorithm, $x = \begin{bmatrix} d \\ z \end{bmatrix} = 0$ is always a feasible solution for the subproblem and master problem. Adding this feasible proposal to X^k (e.g., $X^k = X^k + [\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}]$) and assigning a weight (e.g., $\lambda_0 \geq 0$) to it will not alter the master problem formulation (i.e., $e^{kT}\lambda + \lambda_0 = 1$). This would yield, $e^{kT}\lambda \leq 1$ once λ_0 is eliminated from the formulation and the master problem will always have a feasible solution and artificial variables are not needed anymore in the master problem.

Table 18: Effect of excluding artificial variables from the modified DW algorithm

Modified DW Algorithm	***CPU time (sec)***			***DW iterations***
	Sub	*Master*	*Total*	
No artificial vars.	73.4	11,771.5	11,844.9	161
With artificial vars.	122.9	15,971.6	16,094.5	176

4.3 Numerical Results for a Two Region Canadian Energy Model

In this section, we provide numerical results on the performance of the approximate-PIES DW algorithm by using a realistic two-region energy equilibrium model for Canada from (Fuller and Chung; 2005). Fuller and Chung (2005) illustrated their numerical results using the PIES algorithm (with several PIES steps) to solve for the master problem and the subproblem at each iteration of the DW decomposition. On the other hand, the approximate DW algorithm of section 4.2.3 proposes that instead of solving the exact equilibrium master problem, an integrable approximation of it (e.g., $\tilde{G}(x;\varphi^k)$) can be solved. PIES algorithm also approximates the original mapping of the equilibrium problem by an integrable one[26]. It solves the equilibrium problem iteratively until there is not much change in the solution of two consecutive iterations. The approximate DW algorithm, however, propose to solve it only once. Instead of an exact solution or close to exact solution, even an approximate solution (with one step of PIES) within the DW algorithm is sufficient for convergence, as we illustrate below.

We have tested the approximate DW algorithm on a two-region energy equilibrium model for Canada and provide the results for two cases (PIES with up to 20

[26] Suppose that an inverse demand function is given in the form of $P(d)$ and it is not integrable (i.e., the equilibrium problem cannot be converted into an equivalent economic surplus maximization problem). The integrable approximation to this function is obtained by a modified inverse demand function, $P(d;d^k)$ with its i^{th} component as $P_i(d_1^k,\dots d_{i-1}^k,d_i,d_{i+1}^k \dots,d_n^k)$. PIES algorithm starts by choosing an initial estimate d^0 for the equilibrium solution d^*, and solves the equilibrium problem with the integrable approximation $P(d;d^k)$ by an equivalent nonlinear programming (economic surplus maximization) problem. Setting the obtained solution to d^{k+1}, the algorithm iteratively solves for a series of integrable problems and terminates when $|d^{k+1}-d^k|$ is less than a convergence tolerance. The PIES algorithm, with its approximation at each iteration, has a connection with the well-known nonlinear Jacobi method for solving a system of nonlinear equations (Ahn and Hogan, 1982).

steps and PIES with only one step). The results on progress of the approximate-PIES DW algorithm are summarized in Table 19.

Table 19: Progress of iterations for the approximate-PIES DW algorithm for energy equilibrium model of Fuller and Chung (2005)

iteration	**PIES with 1 Step**				**PIES with up to 20 Steps**			
	CG	***CG/ surplus (%)***	***max(q_m-q*) /q* *100%***		***CG***	***CG/ surplus (%)***	***max(q_m-q*) /q* *100%***	***PIES steps***
1	-13699.661	9040.771	944.245		-11262.152	7432.194	944.242	19
2	-388095.091	256114.280	279.607		-401424.788	264910.902	91.470	16
3	-401323.562	264844.100	470.252		-397991.302	262645.053	78.529	8
4	-171625.821	113260.447	537.951		-499.243	329.463	22.548	18
5	-36803.256	24287.448	282.942		-137204.945	90545.195	58.016	17
6	-250.543	165.340	20.076		-191.454	126.346	9.469	14
7	-40664.193	26835.384	150.304		-7177.857	4736.859	3.242	16
8	-150.977	99.634	53.117		-108.437	71.560	4.207	12
9	-7406.498	4887.745	37.092		-78.011	51.482	7.310	14
10	-36.450	24.054	8.815		-22.505	14.852	1.178	13
11	-37.989	25.070	15.552		-43.594	28.769	6.683	13
12	-92.932	61.328	38.706		-6.073	4.008	1.481	16
13	-17.772	11.728	5.582		-7.261	4.791	1.897	13
14	-8.099	5.345	2.482		-2.084	1.375	0.251	12
15	-10.893	7.188	1.618		-1.282	0.846	0.384	13
16	-4.037	2.664	0.632		-0.616	0.406	0.209	12
17	-4.826	3.185	1.154		-0.202	0.133	0.085	11
18	-1.508	0.995	0.768		-0.218	0.144	0.105	11
19	-1.035	0.683	0.481		-0.222	0.146	0.116	11
20	-1.033	0.682	0.391		-0.059	0.039	0.075	11
21	-0.183	0.121	0.169		-0.014	0.009	0.048	10
22	-0.206	0.136	0.139		-0.011	0.007	0.048	8
23	-0.093	0.062	0.113		-0.006	0.004	0.023	9
24	-0.070	0.046	0.098		-0.004	0.002	0.018	8
25	-0.015	0.010	0.060		-0.003	0.002	0.010	7
26	-0.008	0.005	0.034		-0.001	6.E-04	0.014	6
27	-0.005	0.003	0.032		-6.E-04	4.E-04	0.013	5
28	-0.010	0.006	0.019		-6.E-04	4.E-04	0.012	6
29	-0.001	9.E-04	0.015		-2.E-05	1.E-05	0.002	7
30	-0.002	1.E-03	0.014		-1.E-05	9.E-06	0.002	4

"*CG*" column shows the convergence gap at each decomposition iteration as in Fuller and Chung (2005) and "*CG/surplus(%)*" denotes the relative measure of convergence gap as percent of the producer's surplus at equilibrium solution (e.g., 151.532 billion). Finally, "*max(q_m-q*)/q* 100%*" column is the maximum, over all demand quantities in vector *q*, of the absolute difference between master problem's solution and the reference solution, expressed as a percent of the reference solution. Finally, the last column is the number of PIES steps required[27] at each DW iteration for the DW algorithm with up to 20 PIES steps.

Without any decomposition and using the PIES algorithm for the original model, it takes 0.387 seconds and 9 PIES iterations to reach the equilibrium solution with an accuracy of 0.00635%. The convergence gap as a percent of the producer's surplus for the DW algorithms with up to 20 PIES steps and 1 PIES step after 30 decomposition iterations are 0.000009% and 0.001365%, respectively and both methods have a high degree of accuracy, 9.E-06% and 0.0142%, respectively. However, the computational gain with the approximate DW algorithm (e.g., PIES with 1 step) for this illustration is substantial (79.17% decrease in total CPU time) and this is presented in the next table.

Table 20: Total CPU times for the approximate DW algorithm for energy equilibrium model of Fuller and Chung (2005)

Approximate DW Algorithm	***Master Problem***	***Subproblem for Region 1***	***Subproblem for Region 2***	***Total***
PIES with up to 20 steps	6.574	2.785	3.113	12.473
PIES with 1 step	0.648	0.957	0.992	2.598
% Decrease in CPU time	90.14%	65.64%	68.13%	79.17%

In this illustration, note that both the subproblem and the master problem solutions are approximated in the overall DW algorithm. Chung and Fuller (2010) study approximations of the subproblem for DW decomposition of VI problems and this illustration combines their ideas with the approximation of the master problem. This is important, because any real implementation necessarily has some degree of error in the solutions of the master problem and the subproblem. Moreover, one can control the amount of computational effort required at each iteration in order to decrease the overall computational burden.

4.4 Summary and Discussions

In this chapter, we present DW algorithms for two models in energy markets. The exact DW algorithm fails to converge within a reasonable time for our illustrations.

[27] PIES method with up to 20 steps has the convergence condition that there should be no more than 0.1% change in the price of any demand commodity from one PIES iteration to the next.

However, the modified and approximate DW algorithms converge to the equilibrium solution with significant computational improvements over the exact DW algorithm.

Although the models without any decomposition can be solved considerably quicker than any DW algorithm, the benefits of managing the subproblem and master problem separately may compensate for the additional time to obtain a solution. For the TOU pricing models, separate teams or analysts can maintain the subproblem (containing only firms' problem) and the master problem (containing convex combinations of the proposals from subproblems and the ISO's problem with network constraints). With further approximation in the subproblems (as in Chung and Fuller (2010)), models with special structure can be decomposed by other dimensions. As an example, energy market models can be decomposed by region (e.g., western and eastern Canada) or by commodity (e.g., electricity, gas, oil). Therefore, the resulting subproblems can be managed and maintained separately. A consistent solution of the whole model may be obtained by the proposed DW algorithms.

Moreover, parallel computation of the master problem and the subproblem can increase the computational efficiency. Also, the memory limits may cause problems for a very large-scale problem (e.g., a stochastic model with many scenarios) and the only practical option may be the decomposition of the problem.

Although the numerical results are presented for the VI problems, the theoretical results also hold for variety of problems, e.g., NLPs and monotone nonlinear complementarity problems.

5. Future Research Directions

5.1 Modeling Prospects

In the Ontario electricity market, the move towards the smart metering and new pricing schemes (e.g., TOU pricing) is an important part of the "Integrated Power System Plan" to capture energy conservation opportunities (OPA, 2010). Within this context, design and structure of pricing schemes should deliver the appropriate price signals to consumers to cut energy consumption in on-peak hours. For any pricing scheme, regulatory bodies should ensure that markets are well functioning and market prices reflect sufficiently competitive levels where suppliers are making normal profits (but not monopoly profits).

Market power monitoring and mitigation are key policy issues in the design of the competitive and sustainable electricity markets. Market shares, pivotal supplier measure, concentration indices are typical measures of market power. However, these measures are poor indicators of potential market power because they rely on imprecise measurement of the relevant geographical markets due to the simplification of power transmission network properties (Helman, 2006; Helman and Hobbs, 2010). Welfare analysis based on market price simulations and consumers' surplus approximations can more accurately reflect the potential market power of suppliers. The intention is to propose useful policy tools for market regulators or other market screening entities. It can also provide a quantitative method for analysis of regulatory decisions (e.g., regulated price plan -TOU pricing versus single pricing).

Strategic decisions on network design and capacity substantially affects the short term operational decisions of power suppliers and retailers in the electricity markets. One of the current interests in electricity markets is the integration of long-term investment models (both in generation capacity and transmission network) and their implications on short-term competitive market outcomes using game theoretic approaches. These models concentrate on the efficiency and performance of the wholesale markets as well as the impact of investments on market power issues. The generation and transmission expansion models coupled with operational market models would be suitable for planning and forecasting purposes and regulatory bodies would benefit from these analyses in anticipating and monitoring the strategic behavior of suppliers (either system-wide or at particular locations on the power transmission network). In this context, it is also important to include the stochastic nature of renewable generation (e.g., wind) and to analyze its effects on the market outcomes.

Revenue management and pricing for retailers and suppliers in electricity markets is another line of future research. In the future, subscription based prices (i.e., similar to cellular subscriptions in wireless networks) may be offered by retailers in electricity markets to attract consumers. Due to unique properties of electricity (e.g., non-storability, instant supply-demand balance, uncontrollable flow over network), revenue management among the electricity supply chain agents (from retailers to suppliers) is a challenging problem. Dynamic models of pricing in electricity markets using the vast literature from revenue management and pricing is an interesting and promising research area.

5.2 Computational Prospects

Future research on algorithms and computational efforts for large-scale VI problems include the extension of ideas for modified and approximate DW decomposition algorithms to investigate the Benders decomposition for VI problems. Goffin et al. (1997) and Denault and Goffin (1999, 2005) introduce the analytic center cutting plane method (ACCPM) to solve VI problems. In the context of column generation and cutting planes, ACCPM is a centering concept from interior point methods. A cut is introduced at the center of a feasible region and known to contain the solution. In practice, it is shown to be more effective than other centering schemes. Other cutting methods include Kelley's cutting plane or the one from the dual of DW column generation (i.e., Benders).

ACCPM allows for adding another cut to the master problem along with the cut obtained from the dual information of the subproblem. This cut can be calculated (or approximated) from the analytic center of the feasible region at each iteration of Benders decomposition. It is important to understand whether these cuts should be accumulated in the Benders master problem or they should be updated at each iteration. Intuitively, this issue seems to be related to the convergence characteristics of the Benders master problem (i.e., feasible region of the master problem is monotonically non-increasing, but the convergence gap does not approach zero monotonically). Another research venue is to find conditions under which the convergence properties of the Benders decomposition are maintained by the new ACCPM cuts.

It may also be possible to remove some previously added cuts (associated with the information from subproblems) from the master problem, if these cuts are not binding anymore. Although this approach may lead to improvements in the speed of the algorithm, convergence of the algorithm is uncertain.

Finding the conditions under which the convergence properties of the Benders decomposition are improved by the new ACCPM cuts is also important. It may be possible to remove some previously added cuts (associated with the information from subproblems) from the master problem, if these cuts are not binding anymore. This approach may lead to improvements in the speed of the algorithm compared to the original Benders decomposition.

A direct application of this method to the models of this thesis is that, ACCPM can be used to compute "central prices" (e.g., central prices for congestion based wheeling) at each iteration of the DW decomposition. Instead of solving the master problem at each iteration, central prices can be used to compute a new proposal from the subproblem (e.g., a proposal provided by the ACCPM method). However, computing the analytic center can be computationally very challenging. Therefore, an approximation of the center can be useful within the Benders decomposition of VI problems in order to reduce the computational effort.

As illustrated by the numerical results in section 4.3 (the approximate-PIES DW algorithm for the two-region Canadian energy model), approximation of the subproblem as well as the master problem has significant computational advantages over the exact DW algorithm. Although, our convergence analysis does not support this type of algorithms, the numerical results provide useful insights. This can be studied in the future research as well.

Column generation methods (e.g., DW, simplicial decomposition) usually include some column dropping schemes that drops the columns that are no longer believed to be necessary in order to express an optimal solution (Murphy, 1973; O'Neill, 1977; Patriksson, 1994; Patriksson, 1999; Garcia, 2003; Garcia-Rodenas et al., 2011). The computational aim is to generate profitable columns in the search process of an optimal solution and, hence, to reduce the number of iterations and to increase efficiency of computations. In DW algorithm for VI problems, computational difficulties may arise when solving the master equilibrium problem, because the problem size grows with added columns. But using the background from optimization problems, these difficulties can be alleviated with a column dropping method.

Murphy (1973) provided column dropping procedures for the generalized programming algorithm (i.e., decomposition of a convex nonlinear program such that the master problem is a LP). He shows that under certain conditions all non-basic columns (columns with zero weights) at each iteration can be dropped, except for the initial column retained from iteration 0. The reason for keeping the column retained from iteration 0 is to guarantee that there exists a non-degenerate basic feasible solution to the restricted master problem, and this consequently allows that all optimal solutions to the dual restricted master problem are contained within a compact set. He provides two conditions to drop columns. The first condition is to drop all non-basic columns (except the one in iteration 0), if the basic optimal set of weights (λ^k vector in the master problem 4.3) for all basic variables and slacks associated with these weights for all basic slack variables are greater than or equal to $\epsilon > 0$ (ϵ fixed for all iterations). But this condition can never be used to drop columns again if once it is violated.

The second condition states that starting with an optimal basic solution determined at iteration *k-1* as the trial solution at iteration *k*, if the new column added to the master problem can be pivoted into basis with a weight at least $\epsilon > 0$ (ϵ fixed for all iterations), then all the non-basic columns from iteration *k-1* (except the column in iteration 0) can be dropped. Application of the second condition requires the solution of the succeeding iteration. However, if at any iteration first condition is satisfied (even it has been violated before) and the determinant of the basis solution is greater than some $\delta > 0$ (δ fixed for all iterations), then the second condition is also satisfied at that iteration. Conversely, even if the first condition is violated, it can be reapplied until again violated, if the second condition is satisfied prior to the reapplication of first condition.

The removal of columns with zero weights is a well studied subject in simplicial decomposition. Columns with weights of zero or sufficiently small values in the solution to a restricted problem can be dropped from the restricted master problem (due to the theorems on maximal number of columns to represent an optimal solution). See Patriksson (1994, p.119) and (Patriksson, 1999, ch.9).

In DW decomposition of VI problems, non-basic columns can be determined by the weight vector λ^k (columns with zero weights) at each iteration k and they can be removed from the master equilibrium problem in order to prevent it from getting too large. Another method is to generate several proposals from subproblems before solving the master problem. After iteration among subproblems and master problem starts, the columns with zero weights can be dropped. An important issue is that whether this column dropping technique preserves the convergence properties of the problem or not. Proof of these proposed extensions can be studied in the future research.

References

Acton, J. P., and B. M. Mitchell. 1980. Evaluating time-of-day electricity rates for residential customers. In *Regulated industries and public enterprise: European and United States perspectives.*, eds. B. M. Mitchell, P. R. Kleindorfer. Lexington: Lexington.

Ahn, B., and W. W. Hogan. 1982. On convergence of the PIES algorithm for computing equilibria. *Operations Research* 30 (2): 281-300.

Aigner, D. J. 1984. Welfare econometrics of peak-load pricing for electricity. *Journal of Econometrics* 26 (1-2): 1-15.

Auslender, A., and M. Teboulle. 1999. Lagrangian duality and related multiplier methods for variational inequality problems. *SIAM Journal on Optimization* 10 (4): 1097-115.

Bertsekas, D. P., and J. N. Tsitsiklis. 1989. *Parallel and distributed computation: Numerical methods*. New Jersey: Prentice Hall.

Bhattacharya, K., M.H.J. Bollen, and J.E. Daalder. 2001. *Operation of restructured power systems*. Dordrecht: Kluwer Academic Pub.

Bompard, E., Y. Ma, R. Napoli, and G. Abrate. 2007. The demand elasticity impacts on the strategic bidding behavior of the electricity producers. *Power Systems, IEEE Transactions on* 22 (1): 188-97.

Borenstein, S. 2005. The long-run efficiency of real-time electricity pricing. *Energy Journal* 26 (3): 93-116.

Borenstein, S., and J. Bushnell. 1998. An empirical analysis of the potential for market power in California's electricity industry. *PWP-044r*, University of California Energy Institute, Berkeley, CA.

Borenstein, S., and S. P. Holland. 2005. On the efficiency of competitive electricity markets with time-invariant retail prices. *The Rand Journal of Economics* 36 (3): 469-93.

Borenstein, S., M. Jaske, and A. Rosenfeld. 2002. Dynamic pricing, advanced metering, and demand response in electricity markets. [cited 11/04 2004]. Available from http://www.ucei.berkeley.edu/ucei/PDF/csemwp105.pdf.

Boucher, J., and Y. Smeers. 2001. Alternative models of restructured electricity systems, part I: no market power. *Operations Research* 49 (6): 821–38.

Caves, D. W., L. R. Christensen, P. E. Schoech, and W. Hendricks. 1984. A comparison of different methodologies in a case study of residential time-of-use electricity pricing: Cost-benefit analysis. *Journal of Econometrics* 26 (1-2): 17-34.

Celebi, E. 2005. *Models of efficient consumer pricing schemes in electricity markets*. MaSC thesis, University of Waterloo.

Celebi, E., and J. D. Fuller. 2007. A model for efficient consumer pricing schemes in electricity markets. *Power Systems, IEEE Transactions on* 22 (1): 60-7.

Chen, Y., B. F. Hobbs, S. Leyffer, and T. S. Munson. 2006. Leader-follower equilibria for electric power and NOx allowances markets. *Computational Management Science* 3 (4): 307-30.

Chick, M. 2002. Le tarif vert retrouve: The marginal cost concept and the pricing of electricity in Britain and France, 1945-1970. *The Energy Journal* 23 (1): 97-116.

Chung, W., and J. D. Fuller. 2010. Subproblem approximation in Dantzig-Wolfe decomposition of variational inequality models with an application to a multicommodity economic equilibrium model. *Operations Research* 58 (5): 1318-27.

———. 2005. Dantzig-Wolfe-Jacobi decomposition method for a class of multi-commodity economic equilibrium models. *Working paper*, Department of Management Sciences, University of Waterloo.

Chung, W., J. D. Fuller, and Y. Wu. 2006. A new decomposition method for multiregional economic equilibrium models. *Operations Research* 54 (4): 643.

Chung, W., J. D. Fuller, and Y. J. Wu. 2003. A new demand-supply decomposition method for a class of economic equilibrium models. *Computational Economics* 21 (3): 231-43.

Conejo, A. J., E. Castillo, R. Minguez, and R. Garcia-Bertrand. 2006. *Decomposition techniques in mathematical programming: Engineering and science applications*. Heidelberg: Springer.

Dantzig, G. B., and P. Wolfe. 1961. The decomposition algorithm for linear programs. *Econometrica* 29 (4): 767-78.

Day, C. J., B. F. Hobbs, and J. S. Pang. 2002. Oligopolistic competition in power networks: A conjectured supply function approach. *Power Systems, IEEE Transactions on* 17 (3): 597-607.

Daxhelet, O., and Y. Smeers. 2001. Variational inequality models of restructured electric systems. In *Applications and Algorithms of Complementarity*, eds. M. C. Ferris, O. L. Mangasarian, J. S. Pang. Dordrecht: Kluwer Academic Pub., 85–120.

Denault, M., and J. Goffin. 2005. The analytic-center cutting-plane method for variational inequalities: A quadratic-cut approach. *INFORMS Journal on Computing* 17 (2): 192.

Denault, M., and J. L. Goffin. 1999. On a primal-dual analytic center cutting plane method for variational inequalities. *Computational Optimization and Applications* 12 (1): 127-55.

Dirkse, S. P., and M. C. Ferris. 1996. A pathsearch damped newton method for computing general equilibria. *Annals of Operations Research* 68 (2): 211-32.

Faruqui, A., and S. George. 2005. Quantifying customer response to dynamic pricing. *The Electricity Journal* 18 (4): 53-63.

Faruqui, A., and S. George. 2002. The value of dynamic pricing in mass markets. *The Electricity Journal* 15 (6): 45-55.

Faruqui, A., and J. R. Malko. 1983. The residential demand for electricity by time-of-use: A survey of twelve experiments with peak load pricing. *Energy* 8 (10): 781-95.

Ferris, M. C., S. P. Dirkse, J. H. Jagla, and A. Meeraus. 2009. An extended mathematical programming framework. *Computers & Chemical Engineering* 33 (12): 1973-82.

Ferris, M. C., O. L. Mangasarian, and J. S. Pang. 2001. *Complementarity: Applications, algorithms and extensions*. Dordrecht: Kluwer Academic Pub.

Fuller, J. D., and W. Chung. 2008. Benders decomposition for a class of variational inequalities. *European Journal of Operational Research* 185 (1): 76-91.

———. 2005. Dantzig-Wolfe decomposition of variational inequalities. *Computational Economics* 25 (4): 303-26.

Gabriel, S. A., A. S. Kydes, and P. Whitman. 2001. The national energy modeling system: A large-scale energy-economic equilibrium model. *Operations Research* 49 (1): 14-25.

Gabriel, S. A., and J. D. Fuller. 2010. A Benders decomposition method for solving stochastic complementarity problems with an application in energy. *Computational Economics* 35: 301-29.

Garcia-Rodenas, R., A. Marin, and M. Patriksson. 2011. Column generation algorithms for nonlinear optimization, II: numerical investigations. *Computers and Operations Research* 38 (3): 591-604.

Garcia, R., A. Marin, and M. Patriksson. 2003. Column generation algorithms for nonlinear optimization, I: Convergence analysis. *Optimization* 52 (2): 171-200.

Geoffrion, A. M. 1972. Generalized Benders decomposition. *Journal of Optimization Theory and Applications* 10 (4): 237-60.

Goffin, J. L., P. Marcotte, and D. Zhu. 1997. An analytic center cutting plane method for pseudomonotone variational inequalities. *Operations Research Letters* 20 (1): 1-6.

Green, R. J., and D. M. Newberry. 1992. Competition in the British electricity spot market. *Journal of Political Economy* 100 (5): 929-53.

Han, D., and W. Sun. 2003. New decomposition methods for solving variational inequality problems. *Mathematical and Computer Modelling* 37 (3-4): 405-18.

Harberger, A. C. 1971. Three basic postulates for applied welfare economics: An interpretive essay. *Journal of Economic Literature* 9 (3): 785-97.

Harker, P. T., and J. S. Pang. 1990. Finite-dimensional variational inequality and nonlinear complementarity problems: A survey of theory, algorithms and applications. *Mathematical Programming* 48 (1): 161-220.

He, B., H. Yang, and C. Zhang. 2004. A modified augmented Lagrangian method for a class of monotone variational inequalities. *European Journal of Operational Research* 159 (1): 35-51.

He, B. S., L. Z. Liao, and H. Yang. 1999. Decomposition method for a class of monotone variational inequality problems. *Journal of Optimization Theory and Applications* 103 (3): 603-22.

Helman, U. 2006. Market power monitoring and mitigation in the US wholesale power markets. *Energy* 31 (6-7): 877-904.

———. 2003. *Oligopolistic competition in wholesale electricity markets: large-scale simulation and policy analysis using complementarity models*. PhD dissertation, Johns Hopkins University.

Helman, U., and B. F. Hobbs. 2010. Large-scale market power modeling: Analysis of the US eastern interconnection and regulatory applications. *Power Systems, IEEE Transactions on* 25 (3): 1434-48.

Hobbs, B. F., G. Drayton, E. B. Fisher, and W. Lise. 2008. Improved transmission representations in oligopolistic market models: quadratic losses, phase shifters, and DC lines. *Power Systems, IEEE Transactions on* 23 (3): 1018-29.

Hobbs, B. F., and Helman U. 2004. Complementarity-based equilibrium modeling for electric power markets. In *Modeling prices in competitive electricity markets*, ed. D. W. Bunn, John Wiley & Sons, Inc.

Hobbs, B.F., and J. S. Pang. 2007. Nash-Cournot equilibria in electric power markets with piecewise linear demand functions and joint constraints. *Operations Research*, 55 (1): 113-27.

———. 2004. Spatial oligopolistic equilibria with arbitrage, shared resources, and price function conjectures. *Mathematical Programming* 101 (1): 57-94.

Hobbs, B. F. 2001. Linear complementarity models of Nash-Cournot competition in bilateral and POOLCO power markets. *Power Systems, IEEE Transactions on* 16 (2): 194-202.

Hogan, W. W. 1975. Energy policy models for project independence. *Computers & Operations Research* 2 (3-4): 251-71.

Holland, S. P., and E. T. Mansur. 2006. The short-run effects of time-varying prices in competitive electricity markets. *The Energy Journal* 27 (4): 127-56.

Independent Electricity System Operator (IESO). 2009. The power system. [cited 02/15 2010]. Available from http://www.ieso.ca./imoweb/siteShared/power_system.asp?sid=ic.

———. 2007. Monthly generator disclosure report. [cited 04/04 2007]. Available from http://www.ieso.ca/imoweb/marketdata/genDisclosure.asp.

———. 2005. Market data. [cited 06/15 2005]. Available from http://www.theimo.com. /imoweb/marketdata/marketData.asp.

Joskow, P. L., and J. Tirole. 2004. Retail electricity competition. *NBER Working Paper no. 10473*.

Just, R. E., D. L. Hueth, and A. Schmitz. 2005. *The welfare economics of public policy: A practical approach to project and policy evaluation*. Northampton: Edward Elgar Pub.

King, C. S., and S. Chatterjee. 2003. Predicting California demand response. *Public Utilities Fortnightly* 141 (13): 27-32.

Kirschen, D. S., and G. Strbac. 2004. *Fundamentals of power system economics*. Sussex: John Wiley & Sons, Ltd.

Lawphongpanich, S., and D. W. Hearn. 1990. Benders decomposition for variational inequalities. *Mathematical Programming* 48 (1): 231-47.

Manne, A. S. 1985. On the formulation and solution of economic equilibrium models. *Economic Equilibrium: Model Formulation and Solution*: 1-22.

Mathiesen, L. 1985. Computation of economic equilibria by a sequence of linear complementarity problems. *Economic Equilibrium: Model Formulation and Solution*: 144-62.

Metzler, C., B. F. Hobbs, and J. S. Pang. 2003. Nash-Cournot equilibria in power markets on a linearized DC network with arbitrage: Formulations and properties. *Networks and Spatial Economics* 3 (2): 123-50.

Mitchell, B. M., W. G. Manning, and J. P. Acton. 1978. *Peak-load pricing*. Cambridge: Ballinger.

Murphy, F. H. 1993. Making large-scale models manageable: modeling from an operations management perspective. *Operations Research* 41 (2): 241-52.

———. 1973. Column dropping procedures for the generalized programming algorithm. *Management Science* 19 (11): 1310-21.

Murphy, F. H., and M. V. Mudrageda. 1998. A decomposition approach for a class of economic equilibrium models. *Operations Research* 46 (3): 368-77.

Murphy, F. H., and S. H. Shaw. 1995. The evolution of energy modeling at the federal energy administration and the energy information administration. *Interfaces* 25 (5): 173-93.

Nagurney, A. 1993. *Network economics: A variational inequality approach*. Dordrecht: Kluwer Academic Pub.

O'Neill, R.P. 1977. Technical note - column dropping in the Dantzig-Wolfe convex programming algorithm: computational experience. *Operations Research* 25 (1): 148-55

Ontario Energy Board (OEB). 2007. Ontario energy board smart price pilot. [cited 10/04 2007]. Available from http://www.oeb.gov.on.ca.

———. 2005. Electricity prices are changing: April 1, 2005. [cited 07/18 2005]. Available from http://www.oeb.gov.on.ca./html/en/consumerinformation/2005elecrates.htm.

Patrick, R. H., and F. A. Wolak. 2001. Estimating the customer-level demand for electricity under real-time market prices. *NBER Working Paper no. 8213*.

Patriksson, M. 1999. *Nonlinear programming and variational inequality problems: A unified approach*. Vol. 23. Dordrecht: Kluwer Academic Pub.

———. 1994. *The traffic assignment problem: Models and methods*. Utrecht: VSP Intl Science.

Pollak, R. A. 1990. Distinguished fellow: Houthakker's contributions to economics. *The Journal of Economic Perspectives* 4 (2): 141-56.

———. 1970. Habit formation and dynamic demand functions. *The Journal of Political Economy* 78 (4): 745-63.

Samuelson, P. A. 1952. Spatial price equilibrium and linear programming. *The American Economic Review* 42 (3): 283-303.

Schweppe, F. C., R. D. Tabors, M. C. Caraminis, and R. E. Bohn. 1988. *Spot pricing of electricity*. Boston: Kluwer Academic Pub.

Shahidehpour, M., H. Yamin and Z. Li. 2002. *Market operations in electric power systems: Forecasting, scheduling, and risk management*. New York: John Wiley & Sons, Inc.

Sioshansi, R., and S. Oren. 2007. How good are supply function equilibrium models: An empirical analysis of the ERCOT balancing market. *Journal of Regulatory Economics* 31 (1): 1-35.

Smeers, Y. 1997. Computable equilibrium models and the restructing of the European electricity and gas markets. *The Energy Journal*, 18 (4): 1-31.

Smeers, Y., and J.-Y. Wei. 1997. Spatially oligopolistic model with opportunity cost pricing for transmission capacity reservations –a variational inequality approach. *CORE Discussion Paper 9717*, Universite Catholique de Louvain.

Stoft, S. 2002. *Power system economics: Designing markets for electricity*. Piscataway: IEEE Press.

Takayama, T., and G. G. Judge. 1971. *Spatial and temporal price and allocation models*. Amsterdam: North-Holland Pub. Co.

Taylor, L. D., and H. S. Houthakker. 2009. *Consumer demand in the United States: Prices, income, and consumption behavior*. Third ed. New York: Springer Verlag.

Treinen R. 2005. Shift factors: methodology and example. Presentation, CRR educational class #5, CAISO Market Operations. [cited 08/13 2008] Available from http://www.caiso.com/docs/2004/02/13/200402131609438684.pdf

Varaiya, P., and F. F. Wu. 1997. MinISO: A minimal independent system operator. *Proceedings of 30th Annual Hawaii International Conference on System Sciences*.

Varian, H. R. 1992. *Microeconomic analysis*. Third ed. New York: W. W. Norton & Company, Inc.

Wei, J. Y., and Y. Smeers. 1999. Spatial oligopolistic electricity models with Cournot generators and regulated transmission prices. *Operations Research* 47 (1): 102–12.

Wong, S. 2005. *Alternative electricity market systems for energy and reserves using stochastic optimization*. MaSC thesis, University of Waterloo.

Wu, F. F., and P. Varaiya. 1999. Coordinated multilateral trades for electric power networks: theory and implementation. *International Journal of Electrical Power & Energy Systems* 21 (2): 75-102.

Wu, Y., and J. D. Fuller. 1995. Introduction of geometric, distributed lag demand into energy-process models. *Energy* 20 (7): 647-56.

Yao, J., B. Willems, S. S. Oren, and I. Adler. 2005. Cournot equilibrium in price-capped two-settlement electricity markets. *Proceedings of the 38th Annual Hawaii International Conference on Systems Science*.

APPENDIX A. Approximation of Welfare Change

In this appendix, we show how to approximate the change in consumers' surplus for demand functions with lag effects.

We assume that consumers are deciding on their consumption level for each energy commodity (e.g., off-peak, mid-peak, on-peak electricity), x_i $(i = 1, \dots n)$ at each period t separately and sequentially, i.e., the last period's demands, $\bar{x}_i$, appear as parameters in the current period's utility function. The consumer's problem is to maximize their utility, $U(x_i, y; \bar{x}_i)$, at each period subject to their budget constraints:

$$\max_{x_i, y} U(x_i, y; \bar{x}_i)$$

$$subject\ to \sum_{i=1}^{n} p_i x_i + y = w \qquad (\lambda)$$

where y is the monetary value of all other commodities, w is the budget (income) and λ is the marginal utility of income. The first-order conditions are:

$$\frac{\partial U}{\partial x_i} = \lambda p_i, \ \ i = 1, \dots, n; \qquad \frac{\partial U}{\partial y} = \lambda; \qquad \sum_{i=1}^{n} p_i x_i + y = w.$$

Using these $n + 2$ conditions We can solve for x_i, y and λ as functions of p_i and w (each depending also on parameter $\bar{x}_i$).

We assume that utility now comes from consumption now, and therefore, *extra* utility due to extra $\bar{x}_i$ is zero if there is no consumption now, i.e.

$$\frac{\partial U}{\partial \bar{x}_i} = 0 \text{ , when } x_i = 0.$$

We also assume that the demand functions satisfy the "integrability" conditions[28] with the habit-formation model assumptions (i.e., lagged demand treated as parameters).

[28]A demand system is integrable (i.e., generated by an underlying utility maximization problem) if and only if it satisfies the several conditions detailed in Varian (1992). For example, these conditions are satisfied by five dynamic demand systems, with habit formation, and corresponding utility functions presented in Pollak (1970). Since preferences are unaffected by addition of a constant to a utility function, the further requirement that $\frac{\partial U}{\partial \bar{x}_i} = 0$, when $x_i = 0$ can be achieved for any utility function of Pollak (1970) by adding a constant (independent of this period's consumption but dependent on the previous period's

Now, consider two different price-quantity scenarios, denoted by superscripts A and B. We can approximate the change in utility $U^B - U^A$ by using the average values of partial derivatives evaluated under scenarios A and B as:

$$U^B - U^A \cong \frac{1}{2}\sum_{i=1}^{n}\left(\frac{\partial U^A}{\partial x_i} + \frac{\partial U^B}{\partial x_i}\right)\left(x_i^B - x_i^A\right) + \frac{1}{2}\left(\frac{\partial U^A}{\partial y} + \frac{\partial U^B}{\partial y}\right)\left(y^B - y^A\right) + \frac{1}{2}\sum_{i=1}^{n}\left(\frac{\partial U^A}{\partial \bar{x}_i} + \frac{\partial U^B}{\partial \bar{x}_i}\right)\left(\bar{x}_i^B - \bar{x}_i^A\right)$$

Because of the income constraint the difference term $y^B - y^A$ equals $\sum_{i=1}^{n}\left(p_i^A x_i^A - p_i^B x_i^B\right)$. Using the first order conditions and assuming constant marginal utility of income under different scenarios, $\lambda = \lambda^A = \lambda^B$, we can rewrite the above approximation as:

$$\frac{U^B - U^A}{\lambda} \cong \frac{1}{2}\sum_{i=1}^{n}\left(p_i^A + p_i^B\right)\left(x_i^B - x_i^A\right) - \sum_{i=1}^{n}\left(p_i^B x_i^B - p_i^A x_i^A\right) + \frac{1}{2\lambda}\sum_{i=1}^{n}\left(\frac{\partial U^A}{\partial \bar{x}_i} + \frac{\partial U^B}{\partial \bar{x}_i}\right)\left(\bar{x}_i^B - \bar{x}_i^A\right)$$

The first term is the change in utility as in Harberger's approximation (Harberger, 1971), and the second term is the change in consumer payments. For our lagged demand formulation, the partial derivatives in the last term are evaluated by

---

consumption) to the utility function such that the new utility equals zero when there is zero consumption now. Ideally, to do theoretically justifiable welfare analysis, the present study would start with a system such as in Pollak (1970) (with the added constant), perform the econometric estimation of parameters using data on prices and incomes, and use the estimated demand system instead of (2.4). However, since we are not skilled econometricians, our goals are more modest: we wish only to illustrate how welfare analysis *could* be done on the outcomes of different runs of a TOU equilibrium model. We must rely on estimates of elasticities from the literature (see Celebi and Fuller (2007)), and we have approximated the nonlinear demand functions of Celebi and Fuller (2007) by linear functions, resulting in a non-integrable demand system. Nevertheless, the system is acceptable for our illustrative purposes.

starting with the first order conditions $p_i = \frac{1}{\lambda}\frac{\partial U}{\partial x_i}$ and the inverse demand functions $\boldsymbol{p} = \boldsymbol{B^{-1}}(\boldsymbol{x} - \boldsymbol{a} - \boldsymbol{E}\,\boldsymbol{\bar{x}})$:[29]

$$\frac{1}{\lambda}\frac{\partial U}{\partial x_k} = \boldsymbol{B_k^{-1}}(\boldsymbol{x} - \boldsymbol{a} - \boldsymbol{E}\,\boldsymbol{\bar{x}})$$

Differentiating with respect to $\bar{x}_i$ gives the second-order partial derivative:

$$\frac{1}{\lambda}\frac{\partial^2 U}{\partial \bar{x}_i x_k} = -B_{ki}^{-1}E_{ii}$$

Reversing the order and integrating over this period's demand would give the partial derivative with respect to previous period's demands (assuming $\frac{\partial U}{\partial \bar{x}_i} = 0$, when $x_i = 0$):

$$\frac{1}{\lambda}\frac{\partial U}{\partial \bar{x}_i} = -\sum_{k=1}^{n} B_{ki}^{-1}E_{ii}x_k$$

Therefore, the approximation can be rewritten as

$$\frac{U^B - U^A}{\lambda} \cong \frac{1}{2}\sum_{i=1}^{n}\left(p_i^A + p_i^B\right)\left(x_i^B - x_i^A\right) - \sum_{i=1}^{n}\left(p_i^B x_i^B - p_i^A x_i^A\right) - \frac{1}{2}\sum_{i=1}^{n}\left(\sum_{k=1}^{n} B_{ki}^{-1}E_{ii}\left(x_k^A + x_k^B\right)\right)\left(\bar{x}_i^B - \bar{x}_i^A\right)$$

This last term in the approximation can be defined as the short-term effect of lagged demand (e.g., habit formation) on the consumers' surplus. In the calculation of consumers' surplus changes of section 2.3.2, this last term accounted for 43% to 76% of the whole change in consumers' surplus (in Table 3, Table 4 and Table 6), and it was always a positive contribution to the total, when scenario B is TOU pricing and scenario A is single pricing scheme (the sign of the last term depends on the difference, $\left(\bar{x}_i^B - \bar{x}_i^A\right)$, and it is positive for almost all of the off-peak and mid-peak hours and only negative for on-peak hours, hence the sum over all demand blocks and periods is always positive).

[29] Demand functions are of the form $\boldsymbol{x} = \boldsymbol{a} + \boldsymbol{Bp} + \boldsymbol{E\bar{x}}$ (as in section 2.2.4,(2.4)) and the column vectors and matrices are defined as $\boldsymbol{x} = [x_i]$, $\boldsymbol{a} = [a_i]$, $\boldsymbol{p} = [p_i]$, $\boldsymbol{\bar{x}} = [\bar{x}_i]$, $\boldsymbol{B} = [B_{ik}]$ and $\boldsymbol{E} = [E_{ii}]$. $\boldsymbol{B^{-1}}$ denotes the inverse of $\boldsymbol{B}$, and $\boldsymbol{B_k^{-1}}$ denotes the k^{th} row of $\boldsymbol{B^{-1}}$.

APPENDIX B. PTDF Calculations for Linearized DC Network

In this appendix, we show the PTDF calculations from the non-linear AC power flow equations. It is based on the presentations by Schweppe et al. (1988) and Treinen (2005) with different notation. With the assumption that the reactive power component of the AC power flow equations can be ignored (Schweppe et al., 1988), linearized DC network power flow refers to the real power component of the AC power flow equations. The real power flow from node n to m over line l is:

$$f_l^{nm} = G_l(V_n^2 - V_n V_m \cos(\omega_n - \omega_m)) + B_l V_n V_m \sin(\omega_n - \omega_m)$$

where f_l^{nm} is the amount of real power flowing from node n to node m over line l; ω_n and ω_m are the phase angles; V_n and V_m are the voltage magnitudes at nodes n and m, respectively; $G_l = \frac{R_l}{R_l^2+X_l^2}$; $B_l = \frac{X_l}{R_l^2+X_l^2}$; R_l (resistance) and X_l (reactance) are the line parameters.

Assuming line resistance is negligible (i.e. $R_l \ll X_l$, thus lossless lines) and approximating, $\omega_n - \omega_m \approx 0$ (hence, $\cos(\omega_n - \omega_m) \approx 1$ and $\sin(\omega_n - \omega_m) \approx \omega_n - \omega_m$) and $V = V_n \approx V_m$ with $V = 1$, we can have:

$$f_l^{nm} = B_l(\omega_n - \omega_m)$$

The choice of hub node in the network is arbitrary (but for simplicity a node that has no generation nor consumption is usually preferred) (Hobbs, 2001). The hub node allows for the measurement of all transactions using a single index representing transmission from the hub node (e.g., net injections from transmission lines into node n). This node is designated as k and $\omega_k = 0$.

Let the line-node incidence matrix be denoted by $A = \{a_{\text{ln}}\}$ and $a_{\text{ln}} = 1$ if flow on line l is from node n and $a_{\text{ln}} = -1$ if flow on line l is to node n, and $a_{\text{ln}} = 0$ otherwise. This is a reduced sized line-node incidence matrix (e.g., hub node's column, k, is deleted). Let B be the diagonal matrix with B_l on the diagonal. Then the closed form expression for the PTDF matrix is (Schweppe et al., 1988):

$$PTDF = \text{BA}(\text{A}^\text{T}\text{BA})^{-1}$$

This matrix sets the relation among injections/withdrawals and the flow on lines:

$$F = PTDF\ Y$$

where Y is vector of injections/withdrawals (e.g., y_n) from the hub node to node n, and F is the vector of line flows f_l^{nm}.

The PTDF matrix we get from the above formula has one less column than total number of nodes (*N*), to which we add a zero column to retrieve a size of L (total no. of lines) by *N*.

The linearized DC network load flow equations do not represent any losses (i.e., quadratic resistance losses are possible to represent but not included in this appendix, see Schweppe et al., 1988).

APPENDIX C. Theorems from Fuller and Chung (2005, 2010)

We provide the theorems from Fuller and Chung (2005) in this appendix. Proofs can be found in Fuller and Chung (2005, 2010).

Theorem 1: x^* solves *VI(K,G)* iff there exists $x^* \in R^n$, $\sigma^* \in R_+^m$ and $\omega^* \in R_+^l$ such that all of the following conditions are satisfied:

$G(x^*) - \nabla g^T(x^*)\sigma^* - \nabla h^T(x^*)\omega^* = 0$

$g(x^*) \geq 0,$

$h(x^*) \geq 0$

$\sigma^{*T} g(x^*) = 0$

$\omega^{*T} h(x^*) = 0.$

Theorem 2: x_S^{k+1} solves *Sub-VI*$^{k+1}(\overline{K}, G - \nabla h^T(x_M^k)\omega^k)$ iff there exists $x_S^{k+1} \in R^n$ and $\sigma_S^{k+1} \in R_+^m$ such that all of the following conditions are satisfied:

$G(x_S^{k+1}) - \nabla h^T(x_M^k)\omega^k - \nabla g^T(x_S^{k+1})\sigma_S^{k+1} = 0$

$g(x_S^{k+1}) \geq 0,$

$\sigma_S^{k+1T} g(x_S^{k+1}) = 0$.

Theorem 3: λ^k solves *Master-VI*$^k(\Lambda^k, H^k)$ iff there exists $\lambda^k \in R_+^k$, $\omega^k \in R_+^l$ and $\psi^k \in R$ such that all of the following conditions are satisfied:

$X^{kT} G(X^k \lambda^k) - X^{kT} \nabla h^T(X^k \lambda^k)\omega^k + e^k \psi^k \geq 0$

$h(X^k \lambda^k) \geq 0$

$e^{kT} \lambda^k = 1$

$\lambda^{kT}(X^{kT} G(X^k \lambda^k) - X^{kT} \nabla h^T(X^k \lambda^k)\omega^k + e^k \psi^k) = 0$

$\omega^{kT} h(X^k \lambda^k) = 0.$

Theorem 4: If x_M^k solves *Sub-VI*$^{k+1}(\overline{K}, G - \nabla h^T(x_M^k)\omega^k)$, then x_M^k solves *VI(K,G)*.

Theorem 5: If $CG^k < 0$, then $conv(X^k) \subset conv(X^k)$ (strict inclusion).

Theorem 6: If $CG^k \geq 0$, then x_M^k solves *VI(K,G)* (with assumption 5 holding, e.g., G is strictly monotone).

Theorem 7: In the special case that *VI(K,G)* is a LP, CG^k equals the difference between the value of the dual feasible solution provided by the subproblem and the value of the master problem.

Theorem 8: Either $CG^k \geq 0$ at a finite iteration number k, or $CG^k < 0$ for all k. In the latter case, if G is continuous and the property that any infinite subsequence of $\{(x_M^k, \omega^k, x_S^{k+1})\}_{k=1}^{\infty}$ has at least one limit point is satisfied, then $\lim_{k\to\infty} CG^k = 0$.

Theorem 9: If G is strictly monotone, then the solution to *VI(K,G)* is unique (if the mapping G has the form described in assumption 5, then the solution is unique in d only if $-p(d)$ is strictly monotone; and the solution is unique in d and z only if $c(z)$ is strictly convex).

Theorem 10: If G is strongly monotone and continuous, then either $x_M^k = x_S^{k+1}$ for a finite iteration number k, or $\lim_{k\to\infty}\|x_S^{k+1} - x_M^k\| = 0$ (if the mapping G has the form described in assumption 5, then $\lim_{k\to\infty}\|d_S^{k+1} - d_M^k\| = 0$ only if $-p(d)$ is strongly montone.)

APPENDIX D. Further Numerical Results

In this appendix, we provide further results for the decomposition algorithms, e.g., results on the convergence of the solutions for different algorithms proposed in the thesis.

Table 21: Convergence of Prices and Total Sales for Node #101 in the Modified DW Algorithm (For Perfect Competition TOU Pricing Model as in (3.8) with line limits)

	Prices($/MWh)				Total Sales (MWh)		
iteration no.	Off-peak	Mid-peak	On-peak		Off-peak	Mid-peak	On-peak
1	15.235	27.588	28.600		2,778.72	2,594.61	2,322.85
11	15.340	27.643	28.597		2,773.77	2,593.46	2,323.08
21	15.396	27.727	28.599		2,771.17	2,591.56	2,323.18
31	15.419	27.781	28.601		2,770.19	2,590.32	2,323.20
41	15.437	27.905	28.657		2,769.62	2,587.48	2,322.18
51	15.410	28.070	28.648		2,771.12	2,583.50	2,322.45
61	15.415	28.041	28.644		2,770.83	2,584.18	2,322.53
71	15.396	28.045	28.448		2,771.29	2,583.75	2,326.45
81	15.399	27.990	28.487		2,771.18	2,585.14	2,325.61
91	15.401	27.957	28.540		2,771.12	2,585.99	2,324.53
101	15.401	27.937	28.599		2,771.26	2,586.57	2,323.33
111	15.407	27.927	28.587		2,770.94	2,586.80	2,323.57
121	15.408	27.914	28.600		2,770.86	2,587.15	2,323.30
131	15.404	27.908	28.600		2,771.04	2,587.28	2,323.30
141	15.405	27.907	28.600		2,770.99	2,587.31	2,323.30
151	15.405	27.907	28.600		2,771.03	2,587.31	2,323.30
161	15.402	27.907	28.600		2,771.17	2,587.29	2,323.29
171	15.399	27.906	28.600		2,771.31	2,587.32	2,323.29
177	15.399	27.905	28.600		2,771.31	2,587.32	2,323.29

Table 22: Convergence of Prices and Total Sales for Node #101 in the Approximate DW Algorithm (For Nash-Cournot TOU Pricing Model as in (3.9) with line limits)

iteration no.	Prices($/MWh)				Total Sales (MWh)		
	Off-peak	Mid-peak	On-peak		Off-peak	Mid-peak	On-peak
1	34.149	59.861	67.257		2,015.69	1,942.25	1,704.60
11	33.277	58.424	66.325		2,053.32	1,973.68	1,720.17
21	33.232	58.605	66.827		2,056.80	1,970.06	1,710.71
31	33.211	58.653	66.666		2,057.51	1,968.63	1,713.78
41	33.199	58.726	66.740		2,058.33	1,966.97	1,712.41
51	33.214	58.625	66.394		2,056.78	1,968.89	1,718.93
61	33.169	58.491	66.396		2,058.74	1,972.01	1,718.73
71	33.068	58.546	66.394		2,063.63	1,970.54	1,718.68
81	33.089	58.516	66.424		2,062.64	1,971.32	1,718.13
91	33.025	58.513	66.349		2,065.55	1,971.18	1,719.48
101	33.049	58.555	66.378		2,064.52	1,970.25	1,718.98
111	33.004	58.674	66.489		2,067.04	1,967.51	1,716.90
121	33.061	58.700	66.491		2,064.40	1,966.98	1,716.96
131	33.065	58.681	66.482		2,064.15	1,967.42	1,717.12
141	33.114	58.734	66.514		2,061.93	1,966.30	1,716.61
151	33.141	58.786	66.522		2,060.74	1,965.11	1,716.52
161	33.170	58.768	66.478		2,059.26	1,965.53	1,717.38
171	33.207	58.794	66.405		2,057.34	1,964.84	1,718.85
181	33.233	58.877	66.328		2,056.08	1,962.80	1,720.40
191	33.242	58.939	66.336		2,055.74	1,961.34	1,720.30
201	33.239	58.949	66.338		2,055.91	1,961.10	1,720.26
211	33.260	58.961	66.368		2,054.98	1,960.90	1,719.73
221	33.264	58.968	66.367		2,054.80	1,960.74	1,719.75
231	33.265	58.969	66.365		2,054.76	1,960.72	1,719.79