SOME ECONOMIC ASPECTS OF A RESTRUCTURED ELECTRICITY INDUSTRY

by

Talat Genc

Doctoral Dissertation Submitted to the

DEPARTMENT OF ECONOMICS

In Partial Fulfillment of the Requirements
For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

2003

THE UNIVERSITY OF ARIZONA ®
GRADUATE COLLEGE

As members of the Final Examination Committee, we certify that we have read the dissertation prepared by Talat Genc

entitled Some Economic Aspects of a Restructured Electricity Industry

and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy

Stanley S. Reynolds — Date 10/23/03

Suvrajeet Sen — Date 10/23/03

Ron L. Oaxaca — Date 10/23/03

Date

Date

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copy of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

Dissertation Director, Stanley S. Reynolds — Date 11/14/03

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at The University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library. Brief quotations from this thesis are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the author.

SIGNED: ______________

To The One

ACKNOWLEDGEMENTS

First, I would like to thank my advisors Dr. Reynolds and Dr. Sen. They have great wisdom and personalities in the universe. I will always remember them as "role-models" in my academic and non-academic life. I am indebted to them.

Second, I would also thank Dr. Oaxaca and Dr.Fishback. They were always positive and encouraging during my graduate studies and research.

Finally, I thank my family and friends for having high expectations of me.

TABLE OF CONTENTS

TABLE OF CONTENTS - continued

TABLE OF CONTENTS - continued

TABLE OF CONTENTS - continued

LIST OF FIGURES

LIST OF FIGURES-continued

LIST OF TABLES

ABSTRACT

This thesis examines several issues that arise in restructured electricity markets. These issues include production, scheduling and forward contract decision making, capital investment decisions in uncertain environments, and equilibrium bidding in wholesale electricity auctions.

In the chapter, "Supply Function Equilibria with Pivotal Suppliers", we study the impact of pivotal suppliers in supply function bidding settings. Observers of these markets have noted the important role that pivotal suppliers, those who can substantially raise the market price by unilaterally withholding generation output, sometimes play. However the literature on SFE has not considered the potential impact of pivotal suppliers on equilibrium predictions. This is a potentially important deficiency of applications of SFE to electricity markets, given the large role that pivotal suppliers sometimes play in these markets. We formulate a model in which generation capacity constraints can cause some suppliers to be pivotal. In symmetric and asymmetric versions of the model we show that when pivotal suppliers are present, the set of SFE is reduced relative to when no suppliers are pivotal.

In the chapter, "Dynamic Oligopolistic Games Under Uncertainty: A Stochastic Programming Approach", we study several stochastic programming formulations of dynamic oligopolistic games under uncertainty. It is well known that if the number of state variables increases, dynamic programming becomes computationally intractable. For such games, we show that under certain symmetry assumptions, players earn greater expected profits as demand volatility increases. The key to our approach is the "scenario formulation" of stochastic programming. The examples presented in this paper illustrate that this approach can address dynamic games that are clearly out of reach for dynamic programming, the common approach in the literature on dynamic games.

In the chapter, "Scenario-based Electricity-Gas Forward and Spot Pricing and Load Formulations", we propose load models and price and return formulations in specific energy markets. Existing energy models do not consider inter-relations between the trio: spot price, derivative price and electric load. Also these models, which are in the spirit of the models proposed in financial and commodity markets, ignore special characteristics of electricity, which may make the proposed models useless. In our formulations we consider these characteristics and correlations between these variables.

Simulation results that we run support our modeling approaches.

CHAPTER 1

INTRODUCTION

Robert Pitofsky, Chairman of the Federal Trade Commission (FTC) told the House Judiciary Committee in June 1997 that "... The electric power industry is on the verge of substantial deregulation. While it is unclear whether that process will be driven by the states or by the federal government, the outcome in either case should be that market forces will have an effect on firms long accustomed to the slower pace of regulated life." Also he said, "The potential for consumer savings and increased choice is enormous, but it is certainly not guaranteed." Currently restructuring in the electricity industry is still in progress. In the second chapter we discuss the evolution of this process.

In chapter 3, we study "Scenario-based Electricity-Gas Forward and Spot Pricing and Load Formulations". We propose models that are used in DASH (Decision Aids for Scheduling and Hedging). The DASH model for Power Portfolio Optimization is software that provides a tool, which helps decision-makers coordinate production decisions with trading opportunities in the wholesale power market. Input data (such as load, forward and spot prices of power and fuels) generated from some statistical models are fed into DASH. By means of Stochastic Programming approaches, DASH makes investment decisions. More specifically, at the start of each month, financial analysts/traders for the producer wish to reevaluate/rebalance their power portfolio (in the current version of DASH, the power portfolio is comprised of physical forward contracts). At this point, they may invoke some decision model (e.g. DASH) which recommends the mix of power products that the producer ought to hold. While the decision model itself may be dynamic (as in DASH), the trader only commits to a recommendation for the current month. After the appropriate rebalancing trades are executed, the traders wait and observe the market until the end of the month, at which point, they update the decision model by "rolling the horizon" forward, and providing up-to-date information to the decision model which then provides an updated recommendation for the next month. While it is possible to use the DASH model at decision-epochs that are less than a month long, the portfolios within DASH are represented at monthly intervals. In this chapter, based on data obtained from Pinnacle West Capital, which is a holding company for Arizona Public Service, the largest investor owned electric utility in Arizona, we propose several models describing the behavior of electric demand (load) in Phoenix (AZ) area by using ARIMA and GARCH type of models. In some of these models, we considered load as a function of temperature. To model gas and power forward prices we propose both parametric and non-parametric time series models. We also modeled electric spot prices in two different ways. First after we derived the forward prices, we formulated spot prices based on the forward prices. This approach is commonly used in modeling financial derivatives. Second, we modeled spot prices independently from forward prices but as a function of gas prices. This approach is supported by Eydeland and Geman (1998), in which it is noted that " The non-storability of electricity leads to the breakdown of the relationship which prevails at equilibrium between spot and future prices on stocks, equity indices, currencies, etc. The "no-arbitrage" argument used to establish this relation is not valid in the case of power, since it requires that the underlying instrument be bought at time t and held until the expiration of the futures contract." Through scenarios generated by the above mentioned statistical models and optimization, the stochastic programming model used in DASH selected portfolio positions that performed well. When compared with a commonly used fixed-mix policy, our simulation experiments demonstrate that the DASH model provides significant advantages over several fixed-mix policies.

In chapter 4, we study "Dynamic Oligopolistic Games Under Uncertainty: A Stochastic Programming Approach". In this paper we have studied several game structures played under uncertainty. We include a game with probabilistic scenarios (GPS), a game with expected scenarios (GES), and a hybrid game (HG). These games are different in terms of information structures that players share. Each player strives to maximize his/her own payoff function by committing production and capital investment decisions. Even though these games can be applied to many sectors, our examples in this paper are based on the electricity industry. The contributions of this paper are the following. First, it is well known that as number of state variables increase, the 'curse' of dimensionality afflicts. Hence the commonly used approach –Dynamic Programming- fails to compute the equilibria. In this case we show how to use the so-called 'scenario based' method of Stochastic Programming to solve large-scale oligopolistic games under uncertainty. Second, we show that under certain assumptions the players' expected profit increase as volatility rises. Third, in the stochastic programming literature it is well known that for monopolies replacing the probabilistic scenarios (a scenario is a plan or a vision that might happen in the future) by an expected scenario, i.e. assuming future is deterministic, the expected profits obtained from such expected value models can be shown to be lower than that obtained from a stochastic decision model (see e.g., Birge and Louveaux [1997]). The consequences of these were unknown for a game environment. We show that for oligopolies, players may be made better off by implementing expected scenarios into their decision-making strategy. However by introducing the HG formulation we show that players intend to deviate from GES to GPS. Hence we conclude that, it is plausible to consider the GPS formulation.

In chapter 5, we study "Supply Function Equilibria with Pivotal Suppliers". A 'supply function' is a strategy specifying the quantity that a firm is willing to produce as a function of the market price. Supply function equilibria (SFE) have been studied by Grossman (1981) and Hart (1985) under the absence of uncertainty. However there are two problems in studying SFE in this environment. First, the number of equilibria supported by supply functions is enormous. Second, under certainty the firm knows its equilibrium residual demand for sure. Hence by choosing either a fixed price or a fixed quantity, the firm can optimize its objective function. Thus there is no incentive to implement a supply function strategy. However it is shown in Klemperer and Meyer (1989) that under uncertainty firms are willing to choose a supply function strategy rather than choosing simple price or quantity strategies. Under uncertainty, for each outcome of the random variable, the firm can find a price and a quantity that optimizes its objective function. Hence the supply function maps each optimum level of price onto optimum quantity. This strategy is of course better than committing to a fixed price (Bertrand type) or fixed quantity (Cournot type) strategies under uncertainty. Although under certainty there are an enormous number of equilibria in supply functions, in the uncertain environment, the set of equilibria shrinks. Under certain demand and cost assumptions, a unique supply function equilibrium can be obtained for symmetric oligopolies.

SFE type of strategies are not uncommon in electric industry. Before the demand (electric load) is observed, each firm submits an offer schedule (non-decreasing supply function) specifying the quantity that they are willing to produce as a function of its price. The independent system operator takes these offers and after the real demand is realized, each firm produces at the price, quantity bundle at which its own supply function intersects with its own residual demand curve.

There are several SFE application papers. For example, Green and Newbery (1992) have studied competition in the British electricity spot market. These generators bid supply functions to the grid dispatchers who meets the demand at the lowest cost. They show that at the Nash equilibrium these generators, National Power and PowerGen, that bid supply functions to the grid dispatchers who meets the demand at the lowest cost, make so much profit far above marginal costs and cause deadweight losses. Thus, to increase the competition they suggest a number of firms to be increased although entry takes two to three years and requires significant capital investment. In their analysis, they follow the Klemperer and Meyer (1989) paper set up. Green (1999) studies the electricity contract market in England and Wales. He shows that competition in the contract markets would cause generators to sell much of their power in these markets and hence would result in spot prices (at the Pool) close to marginal production costs. He employs supply function type strategies in the two stage spot market, where there are two suppliers and many buyers. He also allows conjectural variations to model different degrees of competition in the contract markets. He finds that with the Bertrand conjecture (taking other's price fixed), generators will set prices equal to marginal cost. This result is similar to Allaz and Villa's (1993) competitive market outcome. Newbery (1998) studies competition, contracts and entry in the electricity spot markets using analytically tractable models. He employs a supply function type of strategy to model the spot market and a Cournot type strategy to model the contract market. He finds that first, if the number of players (competitors) increases, then the maximum price reached in the pool and the average pool prices are reduced. Second, if the industry has insufficient capacity (i.e. demand shortage covered by contracts), and new investment has a lower marginal cost than existing investment (in this case incumbents would invest), then contracts can deter entry (in the sense that entrants could not offer lower priced contracts). Baldick and Hogan (2002) study capacity constrained supply function equilibria in electricity spot markets. They also consider stability issues of the equilibria and propose a so-called 'function space iteration' method to solve (find) the equilibria numerically. They say that if the firms are asymmetric in capacities and in cost functions then the differential equation approach of solving supply functions may not be effective, because the resulting supply functions may fail the non-decreasing property. Moreover, along with these constraints and according to their definition of stability, they say that many of the proposed possible equilibria are unstable. This instability restricts the range of equilibria and eliminates some equilibria that may be observed in the real life. Hence their solution approach may be useful.

In Texas during June-July 1998, there was a 20,000% price increase observed in the wholesale electricity markets. In 2000-2001 California experienced shocking electricity price jumps and declines. What caused these price spikes? For some, El Niño- a disruption of the ocean-atmosphere system in the tropical Pacific having important consequences for weather around the world- was responsible. For some it was a result of a shortage of supply. For others, a few producers exercised their market power. In the short run, especially in one-day ahead, in hour ahead, and in real time wholesale markets, there is no elasticity of demand. As electricity supply approaches available capacity, the marginal cost of production increases. If peak demand is close to the capacity level, then there are some ways to make huge profits. For example, a producer may be able to cause a supply shortage by not meeting demand, hence because of this shortage, price spikes may occur. Or, they meet the peak demand by charging tremendous prices, if they think that they can put the market into deficit by not meeting the demand. In this second case suppliers are called "pivotal". Since they are able to change the market price dramatically, they have huge market power, even if their market share is not significant. Rothkopf (2001) gives the following example to illustrate the case of a pivotal supplier. "Suppose that there are 18 generators each with a capacity of 5 and one generator with a capacity of 10. Thus, the total capacity is 100, and the largest generator has only a 10% share of the capacity. Suppose that half of the generators have a marginal unit cost of 5 and half of them have a marginal unit cost of 10. Suppose that the large generator has a marginal unit cost of 7. If the supplier act independently and the completely inelastic demand is, say, 75, the price will be 10, the marginal unit cost of the marginal supplier. However, if the inelastic demand is 94, the large generator, by withholding more than 6 units can make the price arbitrarily high..." In this case, largest generator is pivotal.

There are other papers about pivotal suppliers. For example, Bushnell et al (1998) studies a Pivotal Supplier Index (PSI), which measures market power based on a generator's pivotal status. If the capacity of one generator is greater or equal to the residual demand (total market demand minus summation of total maximum capacity of other generators and total exported power) that must be served then this generator can be considered as pivotal supplier, which can exercise market power. For a given period of a time, PSI is a binary variable for a generator such that, if residual demand is greater than 0, PSI equals 1 and the generator is assumed to be pivotal, otherwise it becomes 0 and generator is non-pivotal. Then PSI for a generator is obtained by averaging of PSI's over time, in which this generator is being pivotal. Essentially PSI measures frequency of monopoly power held by a given firm. Another market concentration measure related to the supplier's pivotal status is that Supply Margin Assessment (SMA), is designed by the FERC, is a form of PSI applied to annual peak condition. During the peak hours, if a supplier is pivotal, then this supplier fails the SMA screen test. The next market concentration measure, Residual Supplier Index (RSI), is considered to be more generalized form of PSI, devised by CAISO (see Sheffrin (2001, 2002)). RSI is calculated as the ratio of residual supply (total supply minus largest seller's supply) to the total demand. Here, total supply is the summation of total in-state supply capacity and total net import, total demand is the sum of metered load and purchased ancillary service, largest seller's supply refers to the difference of its capacity and its contract obligation to load. RSI provides significant predictor of price-cost markups. The last index, Supply Margin Hirfindahl-Hirschman Index (Supply Margin HHI), was proposed by Perekhodtsev et al (2002). It is calculated as the ratio of maximum HHI (100*100) to the number of firms in a pivotal group. Hence market power decreases as the number of firms in the pivotal group increases.

In this fifth chapter we have formulated a simple model of a wholesale electricity auction for which the notion of a pivotal supplier has a natural interpretation. As in other SFE models with bounded demand variation, there is a continuum of equilibria. We examine the connection between pivotal suppliers and the set of SFE. In the symmetric and asymmetric versions of the model we show that when pivotal suppliers are present, the set of SFE is reduced relative to when no suppliers are pivotal. The presence of pivotal suppliers eliminates some of the most competitive SFE from the set of equilibria. These SFE's are eliminated even though they do not violate capacity constraints anywhere along the proposed equilibrium path. The extent to which the equilibrium set is reduced depends on observable market characteristics such as the extent of excess capacity, the load factor, the number of suppliers, and the base load capacity factor.

CHAPTER 2

INSTITUTIONAL BACKGROUND OF THE INDUSTRY

Before we discuss antitrust and restructuring issues of electricity industry, let us take a look at some characteristics of the industry.

Characteristics of electricity industry:

The electricity industry has three components: generation, transmission, and distribution. There are many ways to generate electricity: By nuclear fission at nuclear plants, burning fossil fuels (e.g. coal, natural gas, oil) at fuel plants, hydro electric energy at dams, using wind power at the windmills, or by solar energy. As an example we can explain how electricity is generated at nuclear plants: The heat released by nuclear fission (a nuclear reaction type) is used to boil the water in the reservoir. Boiling water produces high-pressure steam, which turns the turbine. The kinetic energy produced by turbine causes the shaft and armature to spin in the generator, in which electric current is generated. Once (high voltage) electricity is generated it is sent by transmission lines to local regions. Here since it is high voltage, first it is transformed into lower voltage, then local distributors send this lower voltage electricity to consumers through their wires. In the US, before the restructuring efforts this trio (generation, transmission, and distribution) is held by vertically integrated companies that are owned by either investors or municipalities, both of which are subject to federal and state regulations.

There are several factors that make electricity different than other commodities. First it is either very costly or impossible to store depending on technologies used for generation. For example, if the technology is nuclear fission, once you break the atoms apart you need batteries to store the energy. This process is extremely costly. Or once you break the atoms you cannot form the original structure from the remaining particles. That is this process is irreversible. If the technology is hydro plant, it is expensive to pump the water back into a reservoir to store the water as potential energy. Second, temperature and humidity factors make the demand (load) highly variable. Increases in these factors drive the demand up, and hence a shortfall of supply is likely. The third factor is the forced and scheduled outages. For example in the Tucson, AZ area, because of the lightning, generators or transmission lines often fail during the monsoon season, causing supply shortages. Keeping these factors in mind, a shortage of electricity can mean blackouts, like San Francisco experienced many times in summer 2000, and/or astronomical prices like San Diego suffered in the same period. Generation in an amount above the transmission capacity is also not good, because transmission lines have limited capacity, this excess electricity may break out the lines and jeopardize the stability of the electricity grid. All these factors and short run demand inelasticity make it harder to match demand and supply, and hence this complicates the operations in the electricity markets.

Antitrust and the restructuring of the US energy industry:

Restructuring attempts in US electric industry go back to the Federal Public Utility Regulatory Policies Act (PURPA) of 1978, which aimed to stimulate co-generation and new entries into electricity generation. The real deregulation process started in both state and federal levels in the early 1990s. However this process is still going on, and there have been many debates regarding how to set up 'suitable' market structures, and how to set up 'efficient' antitrust policies so that competition becomes 'perfect' and customers, and governments benefit from it.

On the other hand in the related sector, the gas industry, deregulation efforts began in 1978 with this Natural Gas Policy Act (NGPA). Prior to that, the price producers charged to pipelines, and the price pipelines charged to local distribution companies were regulated by Federal Power Commission (now Federal Energy Regulatory Commission (FERC)). First, the NGPA started restructuring the production side when they initiated decontrol of wellhead prices. They expected to see increase in production and exploration. However there were still some restrictions on competition and free-entry (note that because of environmental and licensing restrictions, new entry can take three to five years) on this side. As a remedy in 1989 the Natural Gas Wellhead Decontrol Act passed and by the beginning of January 1, 1993 all federal price controls on wellhead prices were abolished. Second, FERC began restructuring the transportation side. They issued Order No 436, which aimed to control pipelines merchant role on transporting their own natural gas. But since this Order could not achieve its goal, FERC issued Order No 636 forcing pipelines to unbundle their 'sales, gathering, and transportation services'. The aim was to minimize pipelines' monopoly power on the sales market. FERC also took some precautions on pipeline grid functions, and forced them to create 'secondary market' in pipeline capacity, where unneeded capacity could be resold. Third, local distribution companies needed to be policed, and state regulators to ensure residential, commercial and industrial customers are not 'hurt' by high prices have done this. Consequently, Ewing et al (2000) say that along with positive achievements in restructuring of generation and distribution segments, mergers and acquisitions increased, and the desired competition level remained unattained. Besides customers still could not choose their natural gas suppliers. They were forced to subscribe to local distribution company, even though American Gas Association says that more than 40 percent of households will be able to choose their own gas suppliers soon.

By FERC Order numbers 888 and 889, issued on April 24th, 1996, the restructuring process in the electric sector began. Before we explain these orders, let us discuss the effects of the PURPA on electricity industry. By the PURPA orders many utilities signed long-term purchase contracts at very high prices. Those prices looked especially bad as the cost of natural gas fell in real terms through the 1980s and 1990s, making other generation sources much less economic. For example, during the same period increased safety regulations, many forced outages at plants, unforeseen construction costs, and increased waste disposal costs made nuclear power expensive. Because of these long term contracts and increased nuclear power costs, some states experienced high electric prices in the 1990s. However other states who did not undertake these contracts and did not use nuclear plants kept low electricity prices.

According to Borenstein and Bushnell (2001) these price discrepancies between the states triggered the restructuring efforts in the US. They also say that the most important gain from the restructuring is that the sector will change the way it makes investment and consumption (retail consumer) decisions. Order No 888 was about transmission services. Initially utilities would allow their transmission services to be used by other electricity producers and distributors, and ultimately these utilities would separate their wholesale electricity marketing from transmission services. Order No 889 complemented Order No 888 by imposing restrictions on vertically integrated utilities (owners of plants and transmission grids) so they could not favor affiliates emerging from divestiture of wholesale business from transportation business. However these remedies would not be enough according to Antitrust Division of the Department of Justice and the Federal Trade Commission (FTC) (see Ewing et al (2000)). So they proposed that utilities should transfer their transmission facilities not to their affiliates but to independent system operators (ISOs). However some scholars criticized this proposition, since ISOs were non-profit organizations, and there would not be enough incentives to operate and invest in facilities in an efficient manner. Hence scholars and firms indicated their preferences for independent transmission companies (transcos). In this case transcos would not own any generation and distribution assets, but would be formed by vertical utilities who sold their all assets to other utilities. The reason behind this preference is that these divestitured firms were already familiar with the business and hence were supposed to improve the system efficiency.

Robert Pitofsky, Chairman of the Federal Trade Commission (FTC) told the House Judiciary Committee in June 1997 that "Deregulation in a number of industries (natural gas, long distance telephoning, airlines, trucking and railroad industries) has proven to be beneficial to customers and the competitive process (reducing prices and increasing choices, quality, and innovation)", and "... The electric power industry is on the verge of substantial deregulation. While it is unclear whether that process will be driven by the states or by the federal government, the outcome in either case should be that market forces will have an effect on firms long accustomed to the slower pace of regulated life." Also he said "The potential for consumer savings and increases choice is enormous, but it is certainly not guaranteed" (see more on FTC Testimony: Electric Power Deregulation at www.ftc.gov). Along with restructuring efforts on generation, and transmission side, all states and the District of Columbia have addressed reforms on retail electric service, and according to the Edison Electric Institute 23 states have adopted and begun implementing retail competition (see Ewing et al (2000)).

As FERC Orders 888 and 889 laid the groundwork for a restructured electricity sector, some states began to deregulate their electricity markets. For example in California, in 1996, the state legislature issued a deregulation bill, called AB 1890. It aimed to impose some regulations on power generators and distributors, because electricity prices were higher than that of other states. According to Borenstein and Bushnell (2000) prices were high because companies made large capacity investment, for example by building expensive nuclear plants, which remained idle and they signed unnecessary power purchase contracts at the behest of regulators. For example, before deregulation, utilities, like Southern California Edison and Pacific Gas and Electric, generated their own power. But they made costly mistakes, with the regulator's approval, in building power plants and signing long-term contracts. Also during the deregulation process increases in wholesale electricity prices ruined these companies two times. Since retail prices were regulated, these price increases could not be passed on to the customers and led to bankruptcies. A deregulated wholesale market, where private, unregulated suppliers would compete to produce power was supposed to eliminate these inefficiencies.

On April 1, 1998, California opened up a market for wholesale electricity. Under regulation the cost of new plants and other 'extra' costs were covered by consumers. In contrast under the deregulation plan these costs would not necessarily be covered. But this would mean bankruptcy for some utilities. As a solution three big companies sold some of their plants to other companies so as to cover much of the 'stranded investments' (sunk costs of past investments). The deregulation plan required utilities to sell off many of their power plants either to new incumbents or to some independent non-profit organizations, and required utilities to disconnect their ties with the distribution business and outlawed long term contracts between distributors and suppliers. Also AB 1890 put in some price caps so that power distributors could not charge consumers astronomical prices. Indeed a sub-optimal level of price cap can make both producers (and distributors) and consumers worse off in the following way: If the price cap is too low, in the short run high- marginal-cost generators would be turned off (this might create a supply-shortage), and in the long run, since capacity investment is irreversible producers would lose. If the price cap is too high, this might give incentive to producers to exploit the consumers by exercising market power. Moreover, these high prices can encourage new entries into the market. But these high prices are artificial, because they are due to market power not due to scarcity. Hence not only new enterprises but also incumbent firms might lose in both the short and long runs. (However, the effects of high price cap can be interpreted differently. One might say; new entry can take three to five years. High price persistence can not endure that long. Hence new entry may not occur. Thus incumbents enjoy price increases up to the cap. Or one could say, even if entry occurs, it may increase the competition and ultimately prices go down).

After the deregulation process the restructured California market had three key characteristics. *i*) CAISO (California independent system operator, is a non-profit organization, is responsible from the operations of transmission grid). *ii*) A hybrid market structure was established in which power was exchanged through bilateral contracts, and in the centralized market where it is bought or sold by bidding for shorter periods (spot transactions) and longer periods (forward and options transactions). However many researchers criticized these market structures. For example, David Salant from NERA Economic Consulting said that "…the California auction was structured as a day ahead market that left utilities in a position of having to purchase all their energy needs in a 24-hour period. Day ahead and/or spot markets create vulnerability and uncertainty since utilities are subject to price spikes when suppliers decide they want to limit the amount of power offered." *iii*) To cover the stranded costs, investor-owned utilities would maintain ownership of their transmission facilities, and continue to supply distribution services. However restructuring could not make those stranded investments disappear.

The political side of the restructuring (deregulation) efforts in the electricity industry may be summarized as follows. Democrats generally favor price controls on wholesale power, most Republicans criticized the price controls on the retail side. With this belief, the Democrat-controlled California's legislation passed the 1996 utility deregulation law that regulated the retail side, fixed prices, froze rates, and deregulated the wholesale market. President Bush told Fox News "California needs to address a flawed law. It's a law where they deregulated one aspect of the market and did not deregulate the other aspect of the market. It's a law based on the premise that energy prices would go down. It was a law that prevented hedging in the marketplace. It is a bad piece of legislation, and California needs to change it". Similarly, Treasury Secretary Paul O'Neill said that "It was lunacy for California to keep retail power prices frozen while wholesale prices are market determined". He also said that "California officials were trying to defeat economics…" (see Kucewicz (2001) more on these news). Utilities initially backed the controlled retail prices, because they were expecting to recover their "stranded costs". However later (during the crisis) they started complaining, and asking a judge to abolish the controls on the retail prices. For example senior vice president Mr. Higgins of Edison International, the parent company of South California Edison sent a letter to *The Wall Street Journal*, saying that "Retail rates in California remain frozen not because of the law, but because the California Public Utilities Commission has so far refused to fulfill its responsibility. Under the law, the rate freeze is supposed to end when there is no longer any uneconomic portion of prior investment in utility generation to be recovered…" he also added "…this law is worsening the crisis". Another effect that triggered the crisis is that utilities were required to buy electricity at the Power Exchange, which was created by the 1996 law, almost-immediate delivery at high prices. At the Power Exchange, the single price auction system was used so that highest bid cleared the market even if the other utilities bid lower than that. That means that utilities paid the top-dollar for electricity although they were willing to supply for less. Moreover natural gas price spikes affected the whole country but rates were higher in the west than the rest of the US. Environmental restrictions were also strict in some parts of California. This added extra costs on the production side. Tremendous growth in the west added extra burden on existing generation units. Finally, all of the above factors triggered the energy crisis in California, and ultimately the deregulation experiment failed in year 2000. Borenstein (2001) says that price responsive demand and long term wholesale contracts should be encouraged by the policy makers to solve the problems. He also gives advice to other countries and states saying that "Those states and countries that have not yet started down this road would be wise to wait and learn more from the experiments that are now occurring in California, New York, Pennsylvania, New England, England and Wales, Norway, and Australia, as well as other locations. The difficulties with the outcomes so far, however, should not be interpreted as a failure of restructuring, but as part of the lurching process toward an electric power industry that is still likely to serve customers better than the approaches of the past".

CHAPTER 3

SCENARIO-BASED ELECTRICITY-GAS FORWARD AND SPOT PRICING AND LOAD FORMULATIONS

1 Introduction

The main purpose of load, spot and forward price models used in the decision model DASH is to help generate a finite number of scenarios which are represented in the form of a scenario tree. A scenario models the evolution of information during the decision process (Birge and Louveaux (1997)). It is important to emphasize that our procedures are a combination of statistical methods and heuristics that maintain tractability of the decision model.

The DASH model for Power Portfolio Optimization is software that provides a tool, which helps decision-makers coordinate production decisions with opportunities in the wholesale power market. Input data (such as load, forward and spot prices of power and fuels) generated from some statistical models are fed into DASH. By means of Stochastic Programming approaches, DASH makes investment decisions. More specifically, at the start of each month, financial analysts/traders for the producer wish to reevaluate/rebalance their power portfolio (in the current version of DASH, the power portfolio is comprised of physical forward contracts). At this point, they may invoke some decision model (e.g. DASH) that recommends the mix of power products that the producer ought to hold. While the decision model itself may be dynamic (as in DASH), the trader only commits to a recommendation for the current month. After the appropriate rebalancing trades are executed, the traders wait and observe the market until the end of the month, at which point, they update the decision model by "rolling the horizon" forward, and providing up-to-date information to the decision model which then provides an updated recommendation for the next month. While it is possible to use the DASH model at decision-epochs that are less than a month long, the portfolios within DASH are represented at monthly intervals.

In a completely deregulated market, the traditional notion of load takes a back-seat to demand-curves relating prices and quantities. However the extent of deregulation is in a state of flux in most states in the U.S. For instance in Arizona, retail tariffs are regulated by the state Corporation Commission and are held constant over long periods of time. Electric utilities are required to serve the "native load" that arises from their customers at regulated retail rates. There are several different demand models that have been studied in conjunction with the current DASH model, including time-series that use temperature as one of the main factors. In more humid climates, we expect that humidity will play an important role as well (Feinberg (2002)).

The wholesale electricity market model used in DASH allows electricity forward contracts, and spot market activity. While prices in the electricity market (especially the spot market) vary on an hourly basis, we have discretized time according to a sixteen hour "on-peak" period, and an eight hour "off-peak" period for each day. For the purposes of our model, forward contracts will be assumed to be "monthly", so that planning for period t refers to some month t in the future. Note that the megawatts committed (bought or sold) to the market in period j influences the total electricity generated during period t, $t>j$. To facilitate profit-making, trading decisions must consider future load projections and generation capacity, both of which are subject to uncertainty. If the decisions for the delivery month (t) could be treated independently of other months, then one could develop a model that could treat each delivery month independently. However, such an assumption might expose the firm to a far greater risk level than might be acceptable. This is because (financial) risk exposure of a firm depends on the mix of instruments in its portfolio at any point in time. Hence, it is not sufficient to simply consider profitability for a delivery month; the collection of forwards held at any point in time is an important determinant of risk exposure.

The current price of any forward contract is usually assumed to be known. However, forward prices for each delivery month will evolve over time until the delivery month. As one might expect, this evolution is uncertain on the decision-making date. In the current version of the DASH model, we use a non-parametric approach in which historical data is used to create a vision for the future (e.g., the next six months). This vision is based on creating a number of scenarios of "returns" (percentage change in prices) which may be revealed in the future. The actual process of developing these scenarios is discussed in the next section.

As with forward contracts, "on-peak" and "off-peak" power have different price trajectories, and are modeled separately. However, there are two important observations in modeling the spot market. The time scale for spot prices can be hourly. In the interest of computational tractability, we treat spot markets on a daily basis, and allow spot prices to fluctuate according to the sixteen-hour "on-peak" and eight-hour "off-peak" periods. Also, the spot prices for each day (*d*) during the month (*t*) can be correlated to the forward prices associated with the scenario (*s*) that unfolds.

2.1 Modeling Electricity Demand

Our load data represents an eleven-year period (1990 –2000) of hourly loads in an APS service area. Since each day is modeled by "on-peak" and "off-peak" segments, we begin by transforming the hourly data into averages for "on" and "off" peak segments. The hours 6 a.m. to 10 p.m. are considered on-peak, and the remaining hours are considered "off peak." In order to give the reader a sense of the load data, Figure 2.1 provides a three year sequence of "on peak" loads. On the figure, y-axis that denotes MWh rescaled for the sake of confidentiality. The "off peak" loads also portray similar cyclical and seasonal trends, and these are confirmed by the Kendal-Tau and Turning Point tests.

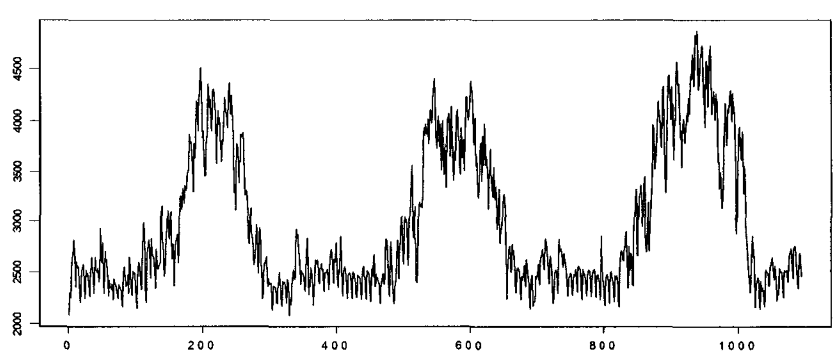


Figure 2.1: On Peak Load Data for 3 Years

Based on seasonality of loads as depicted above, we partition the data for a year into four groups. The first has a decreasing trend, the next an increasing trend and so on. For each group/partition, we use d to denote a day, $\tilde{L}_d$ denotes the load. Assuming an annual growth rate of g we propose the following model:

$$\tilde{L}_d = L_d(1+g) \ ,$$

where $$L_d = \alpha_0 + \alpha_1 d + \varepsilon_d \qquad \text{(Demand Model 1)}$$

$$\varepsilon_d = \Sigma_{i=1}^{7} \beta_i \varepsilon_{d-i} + \eta_d + \beta_8 \eta_{d-1} \ , \quad \eta_d \sim N(0,1)$$

In order to create load scenarios from such a model, we generate standard normal random variates as suggested above.

For the data set we investigated, the de-trended load (for both on peak and off peak segments) followed ARIMA(7,0,1) for each partition. This is consistent with the study of Dupacova, Growe-Kuska and Roemisch (2000) who examined hourly loads (which can be considered as high frequency data) and concluded that SARIMA$(7,0,9)\times(0,1,0)$ was an appropriate model for hourly loads.

In Figure 2.2, we provide plots of the remaining residuals, the autocorrelation, partial autocorrelation functions, p-values of Ljung-Box statistics, and qq-normal plot of residuals. Both ACF and PACF of residuals are within the bounds and Portmanteau test validates this with the qq-normal plots as well. These diagnostics validate the sufficiency of the approach.

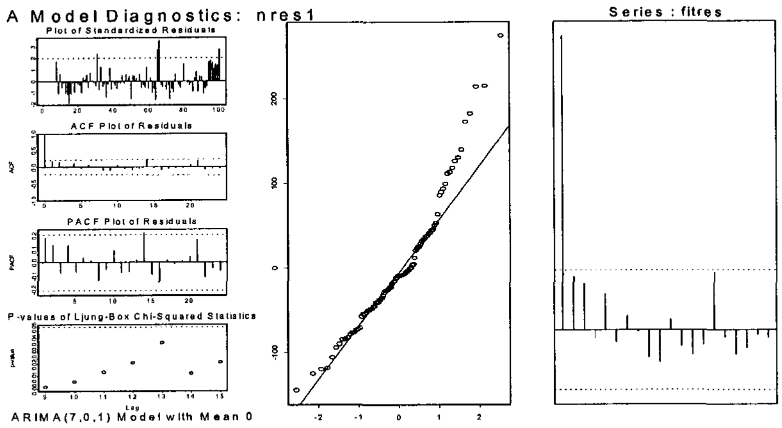


Figure 2.2: Diagnostic Checking of ARIMA(7,0,1) Fitting of De-trended Load Series, Quantile-Quantile Plot, and ACF of Remaining Residuals.

2.1.1 Alternative Load Models

Another load model is based on historical load with drift as well as random noise, which captures cyclicality and seasonality commonly observed in load.

$$L^{y}_{d,s} = \alpha_{0,s} + \alpha_{1,s} L^{y-1}_{d,s} + \varepsilon_{d,s} \quad \text{(Demand Model 2)},$$

where y denotes a year for which the load forecast is being made, and s denotes for a scenario. Given the success of model 1 we used ARIMA (p,d,q) to model the noise. Using ten years of historical load data, once again we find that p=7, and d=0 or 1 and q=1or 7 provide well behaved auto-correlation and partial-autocorrelation functions and Portmanteau test results.

Neither "Demand Model 1" nor "Demand Model 2" use temperature as an explanatory variable. However temperature and humidity are often considered as explanatory variables (Feinberg (2002)). In Arizona (the state for which our forecasts are valid) temperature is a more dominant factor. Using this factor another load model we investigate is as follows.

$$L_{t,s}^{y} = \alpha_{0,s} + \alpha_{1,s} T_{t,s}^{y-1} + \alpha_{2,s} L_{t,s}^{y-1} + \varepsilon_{t,s} \qquad \text{(Demand Model 3)}$$

This load model uses "high" temperature (T) for the previous year to predict on-peak loads and "low" temperature for the previous year to predict off-peak loads. Because of high lag values and persistence of residual load we investigated GARCH (p,q), $p \in [1,5]$,and $q \in [1,2]$. Based on Schwarz's Bayesian criterion (SBC) we have chosen p=q=1 and observed statistically significant coefficients. The autocorrelation function of standardized squared errors suggested serial uncorrelatedness, implying that the model of the error term has captured persistence in the model. For the above model we have also tried ARIMA (p,d,q). Again we obtained the same p, d, q values that we observed in "Demand Model 2."

Decision makers using the DASH model may use any of these alternatives to forecast demand during the study period.

2.2 Modeling Electricity Forward Prices

This part of the DASH model forms the core of our scenario generation procedures. The inputs we use are the forward prices for the preceding year, together with recent trends in the market. Let us first focus on the forward prices for the preceding year. These are available as hourly quotes which we transform into "on-peak" and "off-peak" average prices. We have the following format for the prices: $\pi_{\tau\kappa e}$, where π is the price, and τ, κ and e denote the contract week, delivery week and segment, respectively. Here, the range of indices are: $\tau = 1,2,...52$, $\kappa = \{1,2,...,N\}$, $e \in \{on, off\}$, N denotes the last week in which delivery will happen. For example, $\pi_{1,8,on}$ is the price ($/MWh) on January 7th (i.e. end of week 1) for on-peak power delivered starting on March 1st (for the entire month of March). However we use "returns" to predict prices; that is, $r_{\tau,\kappa,e} = (\pi_{\tau+4,\kappa,e} - \pi_{\tau,\kappa,e}) / \pi_{\tau,\kappa,e}$. Since the index τ reflects a index for weeks, the subscript $\tau + 4$ denotes a period that is four weeks removed from period τ. Assuming that there are four weeks in a month, the return $r_{\tau,\kappa,e}$ denotes the relative change in price during the month starting in week τ.

Note that we use year 2000 forward contract return (prices) to generate 2001 forward contract returns. Given these returns and initial 2001 forward prices we generate 2001 forward curves. For example, to generate "up" and "down" state returns between January and February 2001, we take January 2000 and February 2000 forward contracts . Then we compute the returns from these prices as described above. Then given initial 2001 forward prices (these prices are January prices of the product to be delivered at February, March, April, etc.), we compute February prices of the power to be delivered at March, April, May, etc. Similarly, we use February 2000 and March 2000 forward contracts to generate February 2001 and March 2001 forward curves as described above. Hence, by using this method we take into account of seasonality of forward prices.

There are two important reasons behind this choice (of using returns over prices). First, this approach allows us to treat different power contracts (associated with different months) with the same scenario tree, thus reducing the complexity of modeling the evolution of prices associated with each type of contract. We have empirically verified that it is the interval of time between contract and delivery that is important for modeling returns, and not the actual contract. Hence the same scenario tree remains a valid representation of returns for alternative contacts. Secondly, the econometrics literature recommends that "returns" are better for predictive purposes because empirical evidence suggests that they appear to have better properties (e.g. stationarity) from a computational point of view (Taylor (1986)).

A discrete scenario tree may now be formed by grouping returns into subsets for each period (i.e. month), and modeling the return process as one that allows probabilistic transitions from one subset to another, over time. In order to maintain computational tractability, we consider only two subsets in each period: "High" and "Low" return states. Thus, the resulting scenario tree can be represented by a binary tree in which the returns can assume "High" or "Low" values over the course of the decision process.

To assign "High" and "Low" values for the return states, we adopt a sampling-based procedure that is guided by recent observations of the return series. The nominal value that we assign to each state ("High" or "Low") is the median of the corresponding group for that period. However, without accommodating extreme values, the scenario tree (and consequently the decisions themselves) overlooks extreme events, thus opening up the possibility for catastrophic losses. We will of course, include some loss constraints within the decision model, but in the absence of extreme scenarios, such constraints can only have limited impact. Accordingly, we use a combination of medians and extreme values ("Min" and "Max") to assign values to the High and Low states. The precise manner in which we choose one or the other depends on a heuristic guided by market conditions prior to running the model. Finally, the formation of the scenario tree requires a specification of transition probabilities between nodes which represent information states. Recall that our scenario tree is binary, and hence there are only two probabilities that need to be specified. In the event that our heuristic produced two nodes that are represented by medians (High and Low respectively), then we simply use equal conditional probabilities for these two transitions. On the other hand, if the heuristic produces an extreme value for one path, and a median for the other, then we associate a conditional probability of ¼ for the extreme value, and ¾ for the median value. These conditional probabilities reflect approximately the number of times the return either exceeded the high median or fell below the low-median within the data set. Our heuristic does not produce two extreme values from any node, and hence this possibility is not considered.

The above process creates a binary scenario tree for the return series, which is then used to create prices scenarios that are used within the stochastic programming model described in the following section.

2.3 Modeling Gas Forward Prices

The process used to model gas forward prices is similar to the process described in the previous subsection (on electricity forward prices). We will also assume that the returns for gas and electricity are perfectly correlated (since gas is the marginal fuel) so that a scenario obtained from the electricity forwards returns tree generates a similar scenario from the gas forwards return tree.

2.4 Modeling Electricity Spot Prices

Recall that the forward price process is discretized on a monthly basis. However, spot prices must be modeled on a different time scale. As discussed earlier, on-peak and off-peak spot prices will be modeled on a daily basis, with the understanding that they will be correlated with an appropriate forward price scenario. As with the forwards, we resort to modeling the return series of spot prices.

The spot prices during a delivery month are generated from the following formulation (of spot returns): $r^{p}_{e,d,\tau,\omega} = r^{f}_{e,t,\omega} + \sigma_{e,t}\, z_{d,t,\omega}$, where ω is the node number of the forward scenario tree, $\sigma_{e,t}$ is the standard deviation of spot returns which changes from delivery month to delivery month, and $r^{f}_{e,t,\omega}$ is the daily equivalent of the forward return ($r_{t,t+1,e}$) on node ω for month t. The quantity z represents a standard normal random variate. Here $\sigma_{e,t}$ may be interpreted as the volatility associated with on-peak and off-peak returns during month t and are estimated using a GARCH (Generalized Auto Regressive Conditional Heteroskedasticity) model (Bollerslev (1986)). To validate this estimation method we provide GARCH model outputs on Table 2.1 for on peak spot returns. An appropriate GARCH model has been chosen based on SBC criteria. Because the expectation of spot market prices may be assumed to equal the expected forward prices (Hull (1997)), the above relationship between spot and forward returns captures both the first as well as second moments of the spot price process. Note that the electricity market data for our study reflects prices at Palo Verde, AZ, whereas, the gas market prices reflect data from Henry Hub, LA.

Table 2.1: On peak spot return GARCH model coefficients.

GARCH Estimates of peak returns

SSE	362970.655	Observations	355
MSE	1022	Uncond Var	1552.83427
Log Likelihood	-1682.5152	Total R-Square	.
SBC	3388.51883	AIC	3373.03036
Normality Test	11.4397	Pr > ChiSq	0.0033

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|
Intercept	1	0.6204	1.1868	0.52	0.6011
ARCH0	1	9.0322	5.0182	1.80	0.0719
ARCH1	1	0.0974	0.0219	4.45	<.0001
GARCH1	1	0.8968	0.0216	41.44	<.0001

To help the reader visualize the scenario tree structure, we provide an example in Figure 2.3. In the figure each column denotes a vector of forward prices. Each column is 18 by 1 corresponding to gas and power on and off peak prices. Along the branches we generate spot prices and loads for on-peak and off-peak hours. As an example in figures 2.4 and 2.5, we will provide the first six-month period of 2001 spot prices and load trend projections for on and off peaks based on the scenario defined by arrows on the figure 2.3.

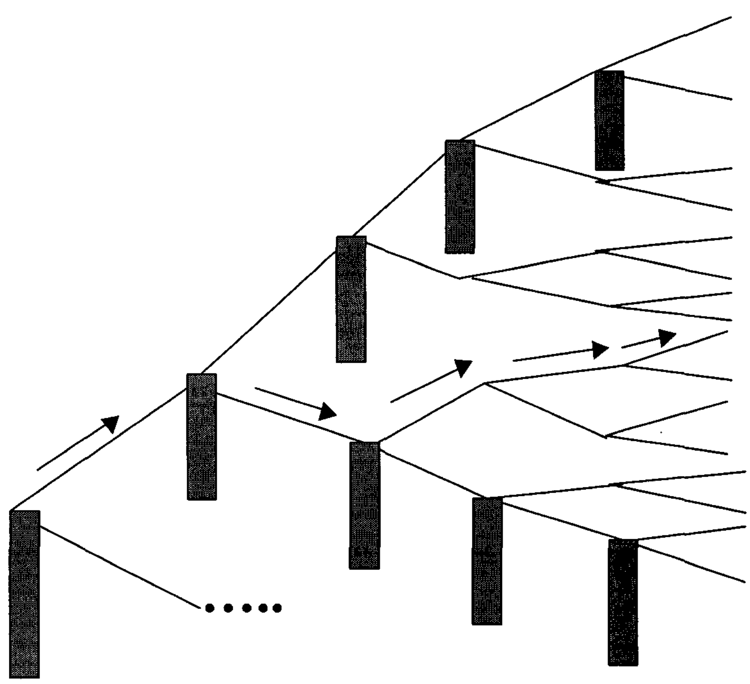


Figure 2.3: Scenario tree structure

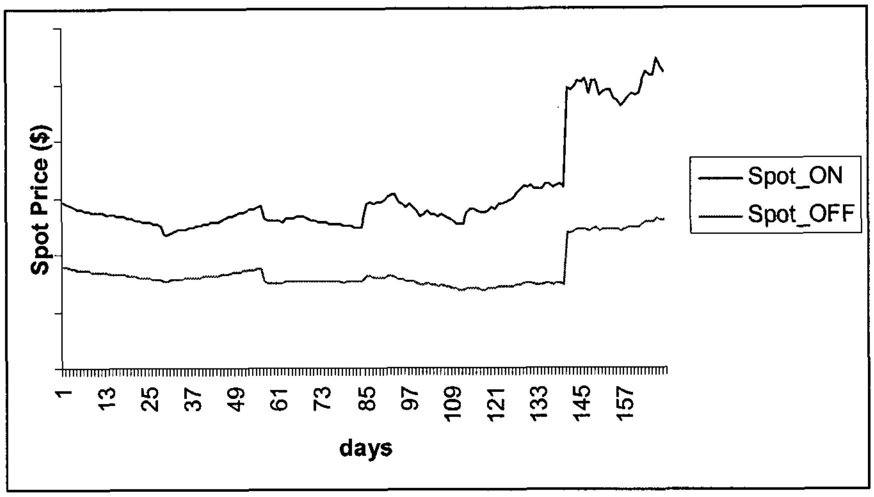


Figure 2.4: On /off peak spot prices for the scenario defined with arrows in Figure 2.3.

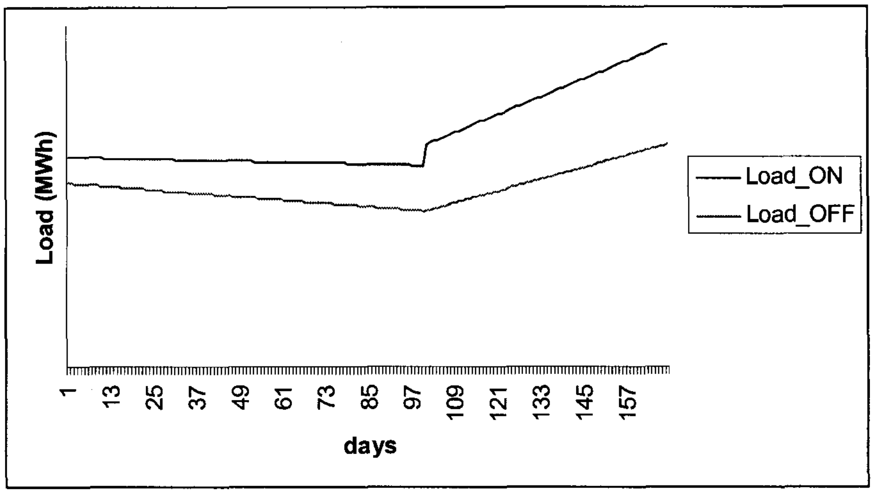


Figure 2.5: On/off peak Load projections for the corresponding scenario.

3 Generating Synthetic Scenarios using Alternate Price Models

So far the above-mentioned models have generated a scenario tree, which is used to create the DASH decisions, was based on first creating forward prices and then generating the other quantities. The results from these decisions were encouraging (see Sen, Yu, Genc (2003)). However, it remains to be seen how the model might perform under realistic scenarios that are not related to the scenario tree used to create the DASH decisions. That is in order to test the robustness of the decisions provided by the DASH model, we modeled the price processes (gas, electricity forwards and spot) directly, rather than modeling "returns" as in formulating the scenario tree. These price models are similar in spirit to studies by Eydeland and Geman (1998), but as one might expect, our prices create scenarios. The details are discussed below, and are significantly different from previous work.

3.1 Gas Spot Pricing

Because gas generators are usually the last ones to be dispatched for merit ordered electricity generation, the marginal cost of electricity reflects gas prices. Hence, it is natural to first model gas spot prices, followed by gas spot price scenarios, and then electricity spot price scenarios. The reader may recall that the DASH model does not allow activities in the gas spot market, and the entire reason for studying gas spot prices is to help generate electricity spot prices. This process of generating gas spot prices uses the following simple algorithm.

Step 0. Calculate trend (μ), and standard deviation (σ) using the first m days of gas spot prices.

Step 1. (Calculate n sample paths for the remaining days). For $i = 1,\ldots M$ (samples) we generate price trajectories for $d = m+1,\ldots,D$ as follows: $\varphi_{i,d} - \varphi_{i,d-1} = \mu + \sigma\varepsilon_{i,d}$.

This process simulates a discrete Brownian motion provided that the trend and volatility remains constant. However, if the forecasting interval (i.e. D) is long, then it may be worth considering a time dependent trend and volatility (e.g. Taylor (1986), Engle (1982)). Dividing the forecasting interval into shorter segments, and then estimating the associated trend and volatility parameters can accomplish this. Our experimental studies were performed for only a five-month interval, and for such a small interval, the estimated parameters were stable over the interval. In Figure 2.6 we provide gas daily spot price scenarios and realized Henry Hub (HH) spot prices using above time independent trend and volatility model. Also in Figure 2.7, we illustrate a comparison of time variant trend and volatility scenarios versus actual HH prices for a short period. From these figures one may conclude that the time variant model may better estimate the realized prices, although the other model seems to be satisfactory.

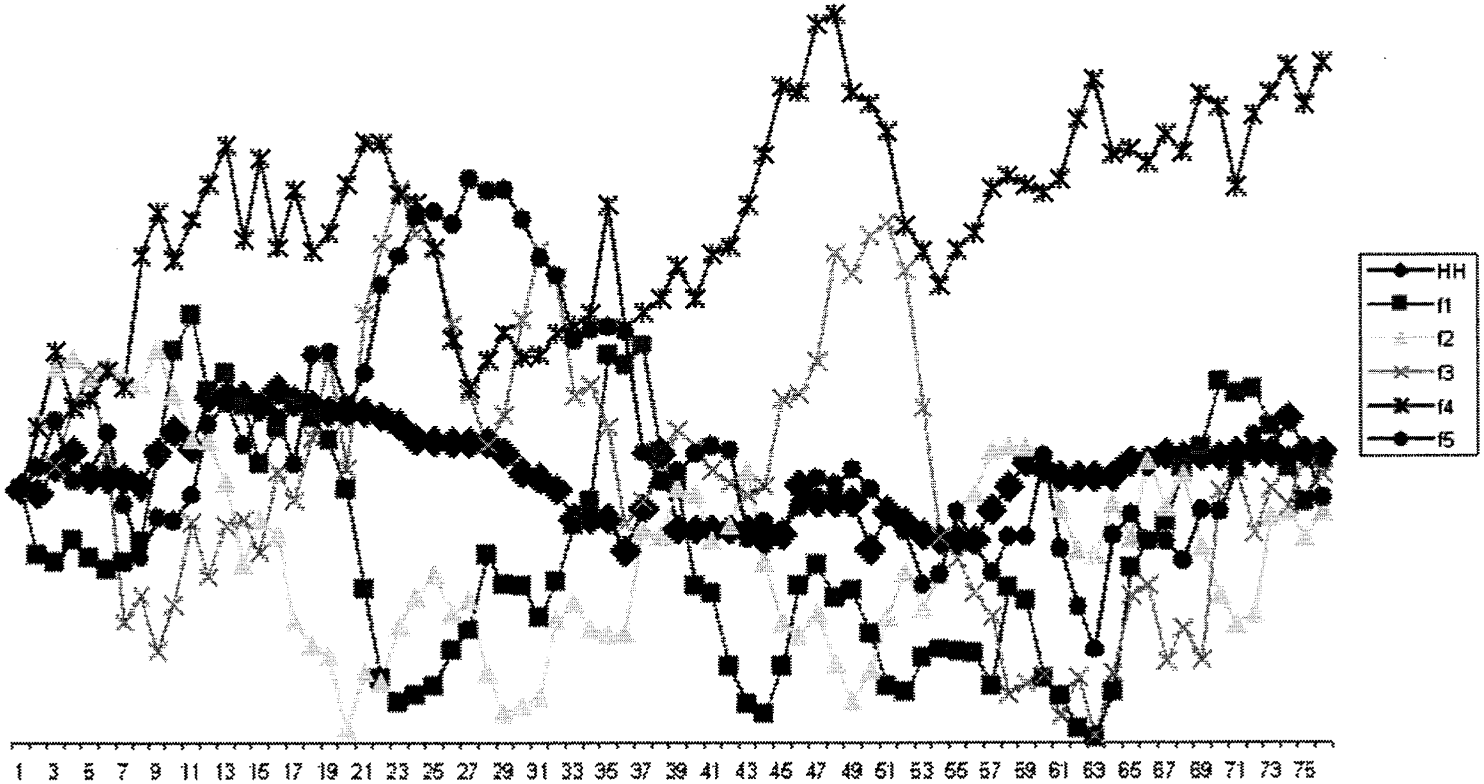


Figure 2.6: Henry Hub prices versus time invariant Brownian motion gas price scenarios.

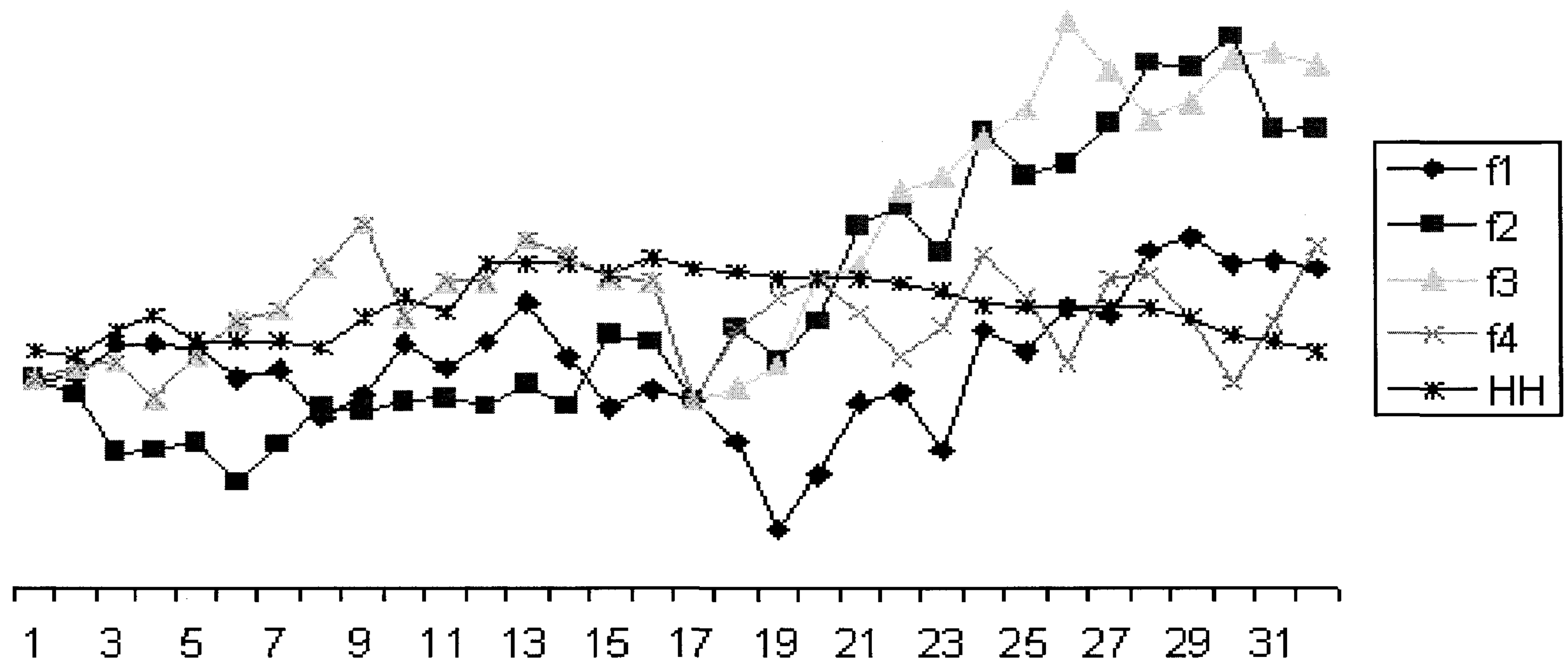


Figure 2.7: Henry Hub prices versus time variant Brownian motion gas price scenarios.

3.2 Spot Pricing of Electricity

In the course of this study, we have observed high correlation between Henry Hub gas spot prices and Palo Verde electricity spot prices. For instance in year 2000 the correlation coefficient between on-peak electricity and gas spot prices is 0.58. The same figure (i.e. correlation coefficient for 2000) for off-peak electricity and gas prices is 0.84. In year 2001 these correlations are 0.62 and 0.7 respectively. As indicated earlier, the reason for this high correlation is that in the Southwest, the marginal fuel is gas. Hence electricity spot prices are very sensitive to gas spot prices in these wholesale markets. Using this information we propose the following relation.

$$(PS)_d^e = \lambda_0 + \lambda_1 (PS)_{d-1}^e + \lambda_2 \varphi_d \xi_d^e .$$

Here, $e \in \{on, off\}$ identifies on-peak and off-peak electricity and the notation $(PS)_d^e, \varphi_d$ denote spot price for on-peak, off-peak electricity and gas respectively for day *d*; moreover ξ_d^e represents a random variable whose distribution we infer from the data.

The motivation for this model is as follows. Since electricity spot prices are persistent, the model includes the lagged spot price and moreover, any change in gas spot prices has a nonlinear effect on the electricity spot price due to the inclusion of the third term in the above equation. Note that while on-peak and off-peak electricity spot prices are correlated, we believe that this correlation is due to their dependence on gas prices, which is reflected in the model stated above. From historical data we have determined the distribution of the error term in the model. For $\lambda_0 = 0$ and $\lambda_1 = \lambda_2 = 1$ we used the "best fit" function of Arena to determine the distribution for the error term. For the on-peak prices the distribution turned out to be the normal with mean 0.1 and 5.5 standard deviation; for off-peak corresponding distribution was Normal with parameters 0.04 and 2. Given initial electricity spot price and gas spot prices generated as in the previous section, we generate electricity spot price trajectories by sampling randomly from the above distributions. Likewise, given other gas spot price trajectories we obtain M on-peak, and off-peak spot price scenarios.

In figures 8 and 9 we compare on/off peak Palo Verde spot prices and scenarios generated by above-mentioned model. From the figures one can observe that some scenarios are good estimates of the realized prices.

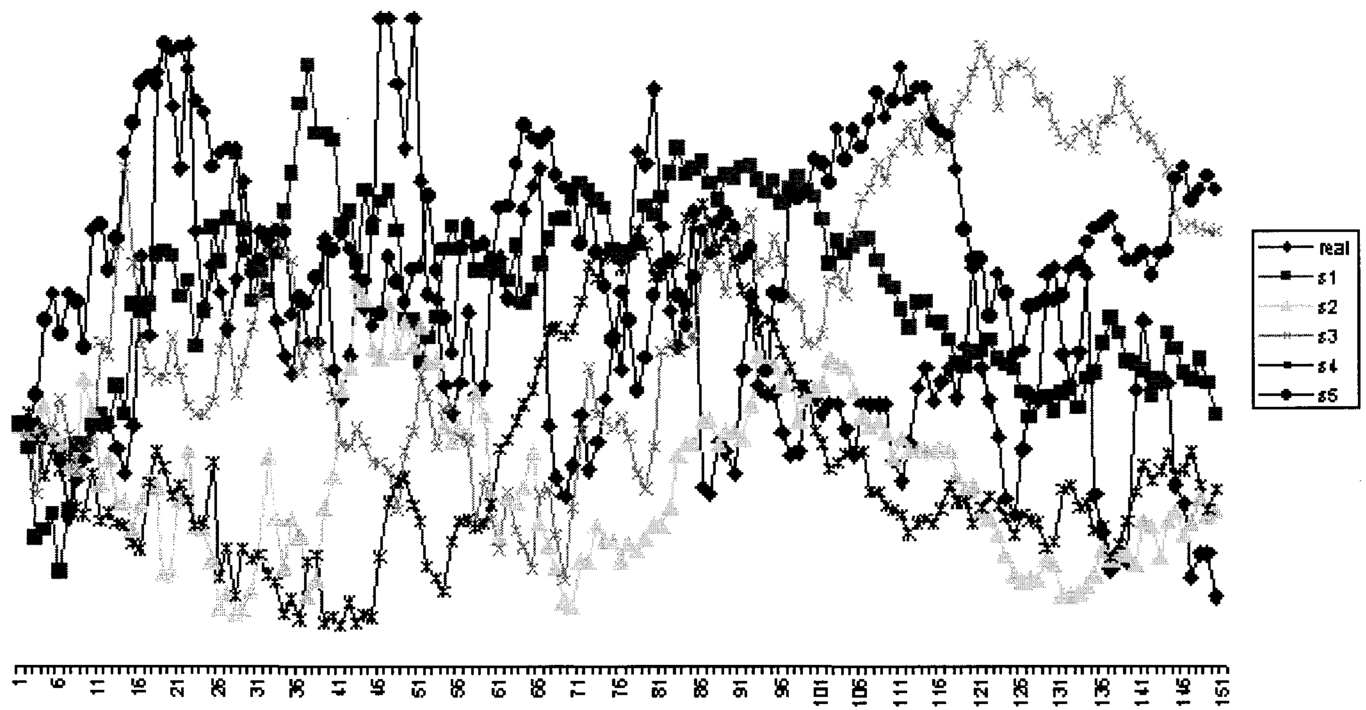


Figure 2.8: Palo Verde electricity on-peak spot prices versus on-peak spot price scenarios.

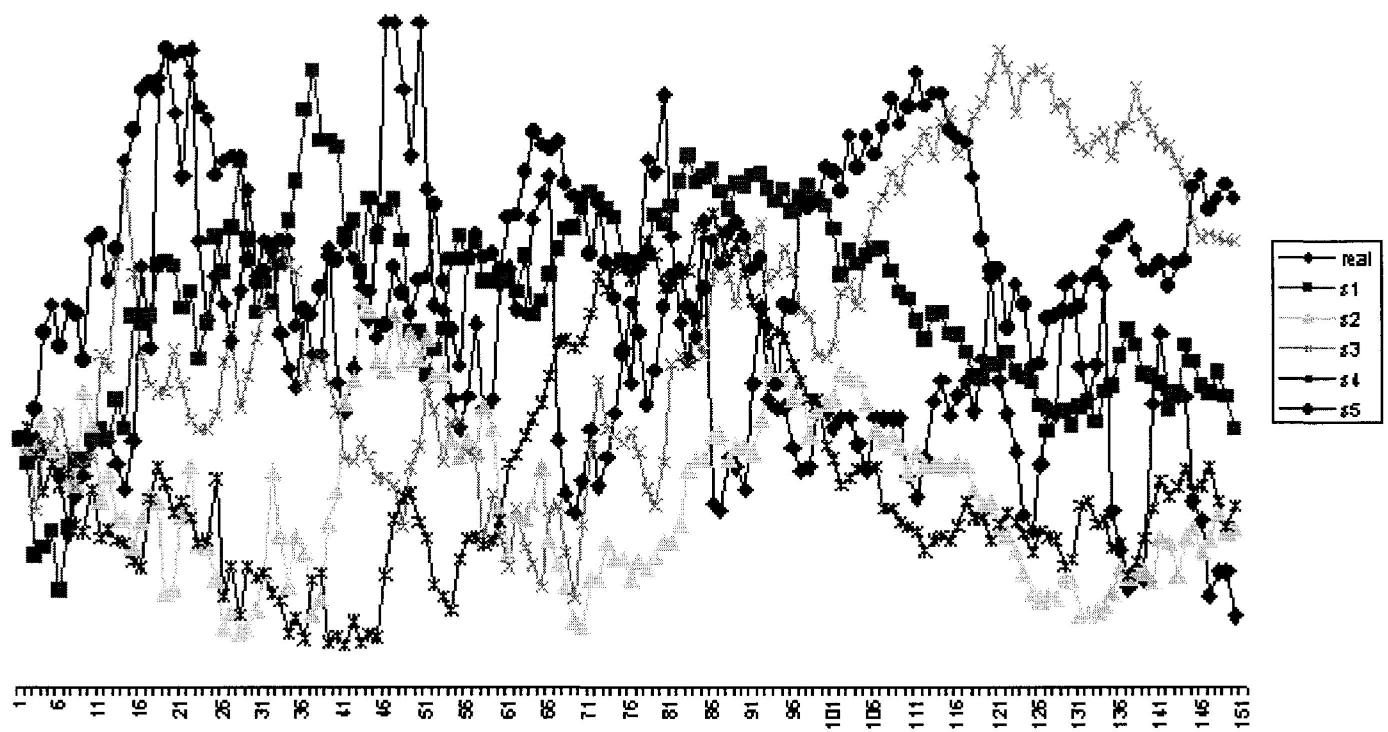


Figure 2.9: Palo Verde off-peak electricity prices versus off-peak spot price scenarios.

3.3 Gas Forward Pricing

We formulate gas forward prices in two steps. First given initial forward prices for five months from now, we determine the trend and standard deviation. Then we use Brownian motion with drift process to generate scenarios. In the following, the unit of time is one month. Specifically,

$GF_{\tau,t} - GF_{\tau,t-1} = \mu + \sigma\varepsilon_t$, $\varepsilon_t \sim N(0,1)$, for $t = 1,2,...,5$, and $\tau \leq t$,

where $GF_{\tau,t}$ denotes the gas forward price at time τ delivered at period t, μ and σ are the drift (trend) term, and standard deviation respectively. We initialize $GF_{0,0}$ as gas spot price at time 0. Given $\tau = 1$ we perform M independent replications of ε_t, for $t = 1,2,...,5$. Then we increment τ and for $t = \tau,..,5$ use M independent replications of $GF_{\tau,t}.(1 + 0.01rand[-l,l])$ to obtain prices in period τ; here *rand* is a uniform random number between $-l$ and $+l$, and l is the estimated standard deviation between prices in period τ and $\tau + 1$.

3.4 Electricity Forward Pricing

It is known that some special characteristics of electricity such as non-storability, weather conditions, unscheduled outages and transmission disruptions make this market different from all other commodity markets. Eydeland and Geman (1998) notes that the market for power derivatives is not complete because hedging portfolios do not exist or are at least very difficult to identify. This incompleteness implies the non-existence of a unique derivative price, hence the wide bid-ask spread observed on certain contracts.

Vehvilainen (2002) concludes that no analytical connection has been established even between the spot price and electricity forward prices and standard financial models may not necessarily apply in electricity markets. Consequently, Eydeland and Geman (1998) propose to approximate power forward prices as

$$PF_{\tau,t} = p_0 + \phi(w_{\tau,t}, L_{\tau,t}),$$

where p_0, $w_{\tau,t}, L_{\tau,t}$, and ϕ designate base load price, forward price of marginal fuel (gas, oil, etc.), expected load (or demand) for date t conditioned on the information available at time τ, and "power stack" function which can either be actual or implied from option prices. Analogous to their formulation, we propose on-peak and off-peak forward prices as

$$PF^e_{\tau,t} = GF_{\tau,t}.\exp(a^e L^e_{\tau,t} + b^e), \qquad (1)$$

where $e \in \{on, off\}$ and a^e, b^e are coefficients. Note that this form of ϕ is rather complex, because load ($L^e_{\tau,t}$) is modeled by either ARIMA or GARCH models, and moreover, the gas forward prices ($GF_{\tau,t}$) are modeled as Brownian motion with some volatility term structure. Note that the standard option pricing models would arise if one were to assume that the load is normally distributed and gas forward prices follow geometric Brownian motion.

In order to implement (1), we follow a sequential procedure that is similar to the one used for gas forwards. Since forward prices are monthly we first convert daily load into average monthly load. Using this load, and gas forward prices we estimate coefficients a^e, b^e by using non-linear least squares. Given $\tau = 1$ we use M previously generated values of $GF_{\tau,t}$ and $L^e_{\tau,t}$ (see sections 4.3 and 2.1 resp.) to obtain *M* values of $PF^e_{\tau,t}$ for $t = 1,2,...,5$. Then we increment τ and for $t = \tau,..,5$ use *M* independent replications of $PF^e_{\tau-1,t}$ to obtain prices $PF^e_{\tau,t}.(1+0.01 rand[-h,h])$ in period τ; here *rand* is a uniform random number between *–h* and *+h,* and *h* is the estimated standard deviation between prices in period τ and $\tau+1$.

4 Conclusions

In this paper we have described how to generate scenario based load, spot and forward price models to be used in decision model DASH. These input models are used in the several experimental tests to study robustness of the DASH. Our experimental evidence suggests (see Sen, Yu, Genc (2003)) that the decision algorithm used in DASH provides a powerful, robust tool for scheduling and hedging in wholesale electricity markets. There are several additional features such as modeling of options, swaps, and forced outages that are being considered as future work to be incorporated into the DASH model.

CHAPTER 4

DYNAMIC OLIGOPOLISTIC GAMES UNDER UNCERTAINTY: A STOCHASTIC PROGRAMMING APPROACH

1. Introduction

Sequential decision-making under uncertainty has a long history of research in a variety of disciplines, including economics, engineering, psychology, and others. Over the past half a century, decision-analytic approaches (e.g., Keeney and Raiffa [1976]) and dynamic programming (e.g., Bertsekas [1987]) have provided the basis for most formal methods for sequential decision-making under uncertainty. Both approaches have been used extensively for policy analysis and modeling. As one might expect, these methods are closely related, although specific details may differ. For instance, decision analysis often focuses on the development of utility and risk assessment, whereas, dynamic programming has been used to develop optimal long-term policies. Despite their numerous applications, these approaches share a common hurdle: the curse of dimensionality (Bellman [1954]). This "curse" is essentially computational, and becomes a burden mainly in the context of computational implementation. Moreover, there are several applications in which the fastest computers are unable to overcome the curse of dimensionality. In order to overcome this difficulty, several researchers have turned their attention to an approach known as Stochastic Programming (SP). In this paper, we formulate multi-player, multi-stage equilibrium problems in which data evolves with the sequential decision-making process, and players make decisions based on a common vision of the future, which is uncertain.

To be sure, the SP approach is not new; it was first proposed by Dantzig [1955], Beale [1955] and Charnes and Cooper [1959] in the context of linear programming under uncertainty. However, the realm of stochastic programming extends far beyond linear programming, and in fact, provides a rich computational framework for decision-making under uncertainty. The power of stochastic programming has remained buried within the optimization literature for a variety of reasons, including the fact that the use of stochastic programming is predicated on the availability of other (auxiliary) optimization software. For instance, the equilibrium problems we consider in this paper require us to use a solver for linear complementarity problems, and such solvers have remained the domain of optimization specialists. With the proliferation of optimization software (e.g., the NEOS web site), this situation is changing, and as a result there has been significant growth in SP applications over the past decade. These include large-scale applications such as financial planning (Carino and Ziemba [1998]), power portfolio optimization (Sen, Yu and Genc [2002]), supply chain management (Fisher et al [1997]), and telecommunications network planning (Sen, Doverspike and Cosares [1994]).

In this paper, we consider stochastic equilibrium problems in which players have a significant stake in technology, and meet their production commitments by investments in a variety of technologies. For example, consider an electric power market consisting of a few players (suppliers), each of whom generates power using a variety of generators.

These generators cannot be installed instantaneously, and as a result, investment precedes production by several months, and sometimes, even years. Under these circumstances, players make their investment decisions under uncertainty. The degree of uncertainty may depend on macroeconomic conditions as well as market-specific characteristics. The SP methodology is based on modeling alternative economic scenarios that may unfold in the future. For the sake of computational tractability, these scenarios are restricted to a finite set, and with each scenario one associates a non-zero probability of occurrence.

The focus of this paper is on the development of models that may be used to predict investment, production, and price trajectories associated with alternative economic scenarios that may unfold. However, these trajectories depend upon the behavior of the players. We will study three alternative behavioral assumptions. In the first formulation, the players make decisions based on collection of probabilistic scenarios, which we refer to as a game with probabilistic scenarios (GPS). Here the trajectories (investment, production, price) will depend on the scenario that unfolds; trajectories will be required to obey a non-clairvoyance condition which states that decisions cannot depend on information revealed in the future. In the SP literature, this condition is also referred to as the non-anticipativity requirement. The second formulation we investigate is called a game with expected scenarios (GES) where investment decisions are based on an expected scenario (as though the world is deterministic). Once the investment decisions are made in a given period, one of the possible scenarios unfolds, and players make their production decisions in response to the specific scenario that unfolds. This type of behavior is not uncommon in some industries where the inclusion of uncertainty within an investment model leads to a very complicated, and sometimes intractable model. For example in the electric power industry, one can invest in a variety of generators (nuclear, hydro, coal, gas etc.) and the resulting capacity expansion models can be rather complex (e.g., WASP-IV [2000], Murphy, Sen and Soyster [1982]).

Due to the difficulties associated with modeling uncertainty within a complex capacity expansion model, players may decide to replace the probabilistic scenarios by an expected scenario, thus leading to a GES game. Nevertheless, we recognize that since operation (generation) decisions are undertaken when better demand information becomes available (i.e. the scenario unfolds), the production game adapts to the scenario that unfolds. Finally, we study a third formulation which we call a hybrid game (HG) which combines features from the GPS and GES games.

Once one moves from the world of single-player optimization to multi-player game environments, the question of the appropriate equilibrium concept arises. There are a variety of types of non-cooperative equilibria for dynamic games. In this paper we focus on the ***S-adapted open-loop equilibrium***. This equilibrium concept was introduced by Haurie, Zaccour and Smeers [1990]. In such an equilibrium each player adopts a strategy that specifies its (production and investment) decisions for each time period and for each possible scenario that can be observed in a time period. In our models, observing a scenario corresponds to knowledge of the current level of demand. A strategy for a firm may be viewed as a contingency plan that specifies actions for each time period and each possible demand state associated with a period. A S-adapted open-loop equilibrium is a set of strategies (plans) for players such that each player's strategy maximizes its expected payoff, given the strategies of the other players.

While the equilibrium concept for GPS is based on that proposed in Haurie, Zaccour and Smeers [1990], our development provides extensions of their work in several directions. First, our analysis points to the fact that competing game models such as GES might seem attractive, but using HG, we argue that the GPS game is the most tenable of the three. In addition, we show that under certain assumptions (i.e., symmetric cost structures), the presence of volatility also provides greater expected profits in a game. This provides the intuition about why players in a market may continue to participate, even though market volatility may be on the rise. In addition, we study multi-stage (sequential) games under uncertainty, and provide a formulation that allows modeling technology additions for applications in which there may be significant lags between the decision to invest, and time at which the plant becomes productive. In contrast to the formulation provided in Haurie, Zaccour and Smeers, we adopt an equivalent scenario-based formulation. The resulting equilibrium conditions are easily applicable for problems with significant lags, and moreover, this formulation is amenable to solution methods for complementarity problems. This helps avoid recursive value function approximations which is the source of the curse of dimensionality in dynamic programming. We illustrate the advantages of this approach with an example that is well out of reach for standard dynamic programming methodology.

An alternative to the open-loop concept is an equilibrium in feedback strategies – the recent economics literature refers to this as a Markov perfect equilibrium (e.g., Ericson and Pakes [1995], Lockwood [1996]). A feedback strategy is conditioned not only on the time period and demand state, but also on a state vector that summarizes the historical decisions that are relevant for players' current and future payoffs.[1] In our model a natural state vector would be the vector of all players' current production capacities. Such a formulation allows one to consider incentives for strategic investment, since the investment decisions of any one player may influence later decisions of other players (see Reynolds [1986, 1987, 1991] for analyses of this type). Note that strategic incentives of this type are absent in the open-loop formulation. A dynamic programming approach is used to solve for a feedback equilibrium.

While the feedback equilibrium approach allows for a somewhat richer set of strategies than the open-loop approach, the feedback approach suffers from some drawbacks. Feedback equilibria are known to exist and be computable only for particular types of environments, such as linear-quadratic settings (see Reynolds [1991] and Lockwood [1996]) or settings with discrete state spaces (see Erickson and Pakes [1995]). Even for models with discrete state spaces, the fact that dynamic programming must be used to compute equilibria means that computations suffer from the curse of dimensionality. To date, equilibria of this type have not been computed for large-scale models. In contrast, open-loop models may be more tractable since dynamic

---

[1] The concept of subgame perfect equilibrium is more general than feedback equilibrium. The set of subgame perfect equilibria may include equilibria in which firms behave collusively; collusion is enforced by threats of reversion to competitive behavior if deviation from collusion is observed.

programming is not necessary to compute open-loop equilibria. We show how to apply stochastic programming methods to solve for open-loop equilibria for large-scale problems. Reynolds [1987] shows that the quantitative difference between feedback equilibria and open-loop equilibria is relatively small for simple dynamic models of production and investment.[2]

We provide conditions for existence and uniqueness of (open-loop) equilibrium for the game formulations that were described above. We also show that players may be worse off in a stochastic game (GPS) than if they play the game deterministically (GES), even though the environment is stochastic. This result distinguishes the multi-player game from a single player setting in which stochastic optimization is known to provide superior expected profits, when compared to deterministic optimization.

The paper is organized as follows. Section 2 introduces the three game formulations for the two-stage case and gives examples which illustrate and compare player's performances and optimal decisions in GPS and GES. Section 3 considers multi-stage models of the games and provides two equivalent formulations which can be characterized as either recursive or non-recursive. Section 4 concludes this paper with a discussion of future research.

2. Some Two-Stage Stochastic Programming Games

We consider a two-period finite player non-cooperative game where players make investment and production decisions under uncertainty about future demand for the

[2] Also, see Deneckere and de Palma [1998] for a defense of using the open-loop concept in dynamic games.

product. In this game, uncertainty can be represented by a finite number of scenarios. Each scenario represents a possible realization of a random process. We assume that in the beginning of the first period all players share a common characterization of the random process. In addition, the players know their own production and investment cost functions and their own initial capacities. Each firm strives to maximize their discounted expected payoff. Given the initial capacity levels, all firms make their initial production decisions (in period 0) to maximize current profit and at the same time they choose their investment in production capacity (in period 0). These investments in capacity become available for production in the next period (period 1). However demand scenarios for the next period are stochastic, and investment decisions must be made here-and-now. After uncertainty unfolds in period one, players make their production decisions. Since this is a two-stage model, we will not consider any investment opportunities in the final period.

2.1 A Two Stage Stochastic Game

Formally this game is defined as follows. Let the players in this game be indexed by $i \in N = \{1, 2, ..., n\}$, where n is finite. In period 0, let $K_i^0 = (K_{i,1}^0, K_{i,2}^0, ..., K_{i,m}^0)$, where $K_{i,k}^0$, k=1,2,…,m, denotes the available capacity at time 0 from technology k for firm i. Since the demand is already realized, the players (firms) choose the production quantities $q_i^*(K_i^0)$ for all $i \in N$ such that when $q_j \equiv q_j^*$, $j \neq i$ then

$$q_i^*(K_i^0) \in \arg\max \{P(Q^*(K^0))\sum_{k=1}^{m} q_{i,k} - c_i(q_i) | \, q_i \in B_i(K_i^0)\} \qquad (1)$$

where $Q^*(K^0) = \sum_{i=1}^{n} \sum_{k=1}^{m} q_{i,k}^*$, $q_i = (q_{i,1}, q_{i,2}, \ldots, q_{i,m})$. $P(Q)$ is the inverse demand function which determines the price of output as a function of total production; this is a strictly decreasing function of total quantity of production in the market. $c_i(q_i)$ is the total cost as a function of the vector of outputs for all technologies[3], and $B_i(.)$ is a convex and compact production set for player i, $B_i \subset \Re_+^m$.

However investment decisions I_i^* must be chosen at period 0 as well, and this decision must be made in the face of uncertainty which is modeled using a discrete random variable whose outcomes represent future economic scenarios indexed by s. Let the investment space for player i be denoted as A_i, which is assumed to be non-empty, and bounded. Thus $I_i^* \in A_i$, such that

$$I_i^* \in \arg\max \left\{ -F_i(I_i) + E[f_i(I_i, q_i^*(K_i^1, \tilde{s}), Q^*(K_i^1, \{K_j^{1*}\}_{j \neq i}; \tilde{s}), \tilde{s}] \mid I_i \in A_i \subseteq \Re_+^m \right\} \tag{2}$$

where period 1 production levels are $q_i^*(K_i^{1*}, s)$ such that,

$$q_i^*(K_i^{1*}, s) \in \arg\max \{ f_i(I_i^*, q_i, Q^*(K^{1*}; s), s) \mid q_i \in B_i(K_i^{1*}) \}. \tag{3}$$

Here $Q^*(K^{1*}; s) = \sum_{i=1}^{n} \sum_{k=1}^{m} q_{i,k}^*(K_i^{1*}, s)$, $f_i(\cdot) = P(Q^*(K^{1*}; s), s) \Sigma_{k=1}^{m} q_{i,k} - c_i(q_i)$, $K_i^{1*} = K_i^0 + I_i^*$. In the above formulation we assume that cost functions $F_i(.)$ *and* $c_i(.)$ are strictly increasing and convex, and $f_i(.)$ is strictly concave in q_i. In addition, it is assumed that the inverse demand functions for periods zero and one are linear in Q, there

[3] More generally, the production cost function would also depend on the amount of capital or capacity. This is essentially a short-run cost of production conditional on the amount of the capital input. For an example of this in a dynamic game analysis, see Reynolds [1986].

exists a Q'>0 such that price is zero for $Q' \geq Q$ (for period one, Q' depends on the scenario) and the price is finite when the quantity is zero.

The strategy space for player *i*, for a given period *t* denoted as Γ_i^t which is compact and convex and the Cartesian product of the production and investment decisions, so that $\Gamma_i^t = \mathrm{A}_i^t \times \mathrm{B}_i^t$, but for the last period $\Gamma_i^t = \mathrm{B}_i^t$, since there is no investment. Given above assumptions and setting there is a unique Nash equilibrium for this two period non-cooperative game (Okuguchi, Szidarovszky [1999]).

2.2 A Two-Stage Game with Probabilistic Scenarios (GPS)

The game formulated in (1), (2), and (3) can also be represented by a model using the so-called scenario formulation. Such a formulation provides a convenient mechanism for games in which players face a large number of constraints. For a given scenario $s \in S = \{1,2,3,\ldots,\omega\}$, $\omega < \infty$, player *i* chooses beginning production quantities as $q_i^*(K_{i,s}^0)$ such that when $q_{j,s} \equiv q_{j,s}^*$, $j \neq i$ then

$$q_i^*(K_{i,s}^0) \in \arg\max \{P(Q^*(K_s^0))q_{i,s} - c_i(q_{i,s}) \mid q_{i,s} \in B_i(K_{i,s}^0)\}. \qquad (1')$$

Here $Q^*(K_s^0) = \sum_{i=1}^n \sum_{k=1}^m q_{i,k}^*(K_{i,s,k}^0)$, $q_{i,s} = (q_{i,s,1}, q_{i,s,2}, \ldots, q_{i,s,m})$.

For each player *i* and scenario *s* investment decisions are made in the first period according to the following.

$$I_{i,s}^{*} \in \arg\max\left\{-F_i(I_{i,s})+[f_i(I_i,q_i^{*}(K_{i,s}^{1},s),Q^{*}(K_{i,s}^{1},\{K_{j,s}^{1*}\}_{j\neq i};s),s] \mid I_{i,s}\in A_{i,s}\subseteq \Re_{+}^{m}\right\}, \quad (2')$$

where period 1 production levels are $q_i^{*}(K_{i,s}^{1*},s)$ which satisfy

$$q_i^{*}(K_{i,s}^{1*},s)\in \arg\max\{f_i(I_i^{*},q_i,Q^{*}(K_s^{1^{*}};s),s) \mid q_{i,s}\in B_i(K_{i,s}^{1*})\}. \quad (3')$$

Again $Q^{*}(K_s^{1*};s)=\sum_{i=1}^{n}\sum_{k=1}^{m} q_{i,s,k}^{*}(K_{i,s}^{1*},s)$, $f_i(.)=P(Q^{*}(K_s^{1*};s),s)\sum_{k=1}^{m} q_{i,s,k}-c_i(q_{i,s})$, $K_{i,s}^{1*}=K_{i,s}^{0}+I_{i,s}^{*}$.

Unfortunately decisions from this formulation cannot be implemented because investment decisions are dependent on future scenarios that are not revealed in the beginning (i.e., period 0). In order to ensure that the decisions can be implemented prior to resolving the uncertainty, we impose the non-anticipativity (non-clairvoyance) constraint as

$$E(I_{i,s}^{*})=I_i^{*} \text{ for all } i \text{ and } s.$$

This equation means that under uncertainty, planning decisions must be implemented before an outcome of the random variable is observed. For illustrative purposes the following example will be confined to a small data set. However the concepts underlying this example are scalable in the sense that many more technologies and decisions can be accommodated within the computational approach of stochastic programming.

***Example 1*:** We develop this example based on observations of wholesale electric power markets in the U.S. Although the data we use in this example is fictitious, it is indicative of the number of players, as well as the technologies used by them in a typical wholesale electricity market. In 1999 generation of electricity in California was based on three major sources of energy: gas, nuclear, hydroelectric, and their percentage shares were 47.2, 17.4, 21.1 respectively. As for the number of players in the market, there existed a total of forty-seven firms in the year 2000, although the five largest producers accounted for over 85% of the production (Pacific Gas and Electric Co (32.58%), Southern California Edison Co (33.29%), Los Angeles Dept of Water and Power (9.54%), San Diego Gas and Electric Co (6.82%), and Sacramento Municipal Utility Distribution (4.34%)). Similarly in Arizona, although there were forty-five producers in the year 2000, the three largest producers accounted for over 85% of the production (Arizona Public Service Co (36.9%), Salt River Project (36.46%), and Tucson Electric Power Co (13.42%)). The 1999 production of electricity may be attributed to the following sources of energy: coal (45.6%), nuclear (36.2%), hydroelectric (12%), gas (6.1%). Finally, the state of New York had sixty-one producers in the year 2000, but the five largest accounted for about 85% of the production (Niagara Mohawk Power Corp (24.32%), Consolidated Edison Co-NY Inc (25.64%), Long Island Power Authority (14.26%), Power Authority of State of NY (11.45%), New York State Elec. & Gas Corp (9.78%). As with the other two states mentioned above, the 1999 generation of electricity in New York was attributed to the following sources: gas (32.1%), nuclear (25.6%), hydroelectric (16.3%), coal (14.8%). From this discussion, it is clear that a market model with three to five players, each using about five technologies may be sufficient to capture most of the production activity in a market, and accordingly, our example will reflect markets of this size.

In this example the production cost function represents the variable generation cost function of a firm. Typical firms own several generation units, including several technologies such as coal, oil, natural gas, nuclear, and hydropower. Since gas, oil fired and coal generators have increasing marginal costs over their operating range of production, we employ quadratic production cost functions. It is clear that the assumption of affine marginal costs does not capture jumps in marginal cost from these technologies and does not capture the rapid increase in marginal costs as output approaches the maximum capacity. However it does represent the qualitative observation of increasing marginal cost with output. For other generation units such as nuclear and hydropower it makes sense to assume constant marginal production costs.

Consider a four-player market with ten available technologies for each player. The data in Table 3.1 shows each player's initial capacities and investment and production cost coefficients from technologies 1 through 10 respectively.

Table 3.1: Cost coefficients and initial capacities from each technology

	Invest./Product. Cost Coeff.	Initial Capacities
Player 1	(6,6,5,5,4,4,3,3,3,3)	(1.5,1.5, 1,0, 0, 1,1,1,0,0)
Player 2	(5,5,5,5,4,4,3,3,3,3)	(1.5,1.5, 1,0, 0, 1,1,1,0,0)
Player 3	(4,4,5,5,4,4,3,3,3,3)	(1.5,1.5, 1,0, 0, 1,1,1,0,0)
Player 4	(3,3,5,5,4,4,3,3,3,3)	(1.5,1.5, 1,0, 0, 1,1,1,0,0)

To interpret this table, initially each player has 1.5 units of the technology 1 and technology 2 capacity, 1 unit of technologies 3, 6, 7 and 8 capacity, and 0 units of technologies 4, 5, 9 and 10 capacity. Investment and production cost functions are $F_{i,1}(.) = C_{i,1}I_{i,1}$ *and* $c_{i,1}(.) = C_{i,1}q_{i,1}^2$, and $F_{i,2}(.) = C_{i,2}I_{i,2}^2$ *and* $c_{i,2}(.) = C_{i,2}q_{i,2}$, with

$C_{i,m} = 7 - i$, for technologies 1 and 2 (i.e. $m = 1,2$) respectively. For technologies 3, 5, 7, and 9 , investment and cost functions are in the following form:

$F_{i,m}(.) = C_{i,m} I_{i,m}$ *and* $c_{i,m}(.) = C_{i,m} q_{i,m}^2$, with $C_{i,3} = 5$, $C_{i,5} = 4$, and $C_{i,7} = C_{i,9} = 3$. And the cost functions of other technologies are defined as $F_{i,m}(.) = C_{i,m} I_{i,m}^2$ *and* $c_{i,m}(.) = C_{i,m} q_{i,m}$, with $C_{i,4} = 5$, $C_{i,6} = 4$, and $C_{i,8} = C_{i,10} = 3$. We may encounter these kinds of technologies in many industries, where there are tradeoffs between investment levels and production costs. In production of electricity, generators such as hydro and nuclear plants are generally expensive to build but inexpensive to operate. In contrast, gas generators require lower capital outlays but may be considerably more expensive to operate.

Each player i (=1,2,3,4) predicts a market with the following scenarios.

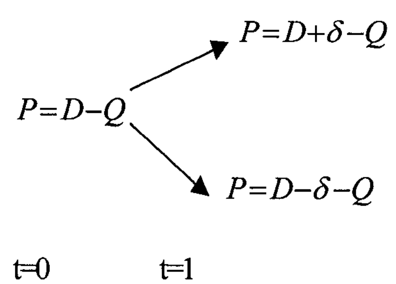


Figure 3.1: Two-stage demand scenarios

where D=90 , δ =5, and u=0.5 is the probability of the "up" state, and d=0.5 is the probability of the "down" state. In the subsequent section we will investigate multi-stage formulations in which a binary tree of this type can be used to approximate a Brownian motion. Thus a simple scenario tree such as the one shown above may be looked upon as a building block for more realistic models.

For player i the expected profit function is the following

$$\Pi_i(.) = P_0 . \sum_k q_{i,k}^0 - \sum_k c_{i,k}(q_{i,k}^0) - \sum_k F_{i,k}(I_{i,k}) + u[P_{up}^1 . \sum_k q_{i,up,k}^1 - \sum_k c_{i,k}(q_{i,up,k}^1)] + d[P_{down}^1 . \sum_k q_{i,down,k}^1 - \sum_k c_{i,k}(q_{i,down,k}^1)] , \quad (4)$$

where P_0 is the initial price and P_{up}^1 and P_{down}^1 are prices in the up/down states respectively in period 1. Then player i's optimization problem is;

$$\max \ \Pi_i(q_{i,s}^t, I_i^t, K_{i,s}^t) \quad (4.1)$$

subject to

$$q_{i,s,k}^t - K_{i,s,k}^t \leq 0 \qquad t = 0,1, \ \forall \ i,s,k \qquad (4.2)$$

$$Q_s^t - \sum_{i,k} q_{i,s,k}^t = 0 \qquad t = 0,1 , \ \forall i,s,k \qquad (4.3)$$

$$q_{i,s,k}^t \geq 0 , \ K_{i,s,k}^t \geq 0 \qquad t = 0,1 , \ \forall i,s,k \qquad (4.4)$$

$$I_{i,s,k}^t \geq 0 \qquad t = 0 , \ \forall i,s,k \qquad (4.5)$$

$$I_{i,k}^t - E(I_{i,s,k}^t) = 0 \qquad t = 0 , \ \forall i,s,k \qquad (4.6)$$

$$K_{i,s,k}^{t+1} - K_{i,s,k}^t - I_{i,s,k}^t = 0 \qquad t = 0 , \ \forall i,s,k \qquad (4.7) ,$$

where $s \in \{up, down\}$.

A few remarks regarding this problem are in order. The objective function in this example maximizes the profit from the initial period plus expected profits of the future. Note that the equilibrium conditions are imposed for every scenario and every period t because of (4.3). Also, (4.6) enforces the non-anticipativity constraint discussed earlier. After deriving Karush-Kuhn-Tucker (KKT) conditions for the above problem we obtain a linear-complementarity problem, which are the natural form of optimality conditions in inequality-constrained problems. We employed Argonne National Lab's NEOS server-

PATH solver which is a well known robust and efficient solver for such linear complementarity problems. For each technology we report each player's optimal investment and production decisions in Table 3.2.

Table 3.2: Equilibrium Production and Investment by Technology

Technology 1:

	Production(t=0)	Investment(t=0)	Production(t=1,up)	Production(t=1,down)
Player 1	1.5	0	1.31191	0.775366
Player 2	1.5	0	1.5	0.915497
Player 3	1.5	0	1.5	1.11058
Player 4	1.5	0.0342923	1.53429	1.41611

Technology 2:

	Production(t=0)	Investment(t=0)	Production(t=1,up)	Production(t=1,down)
Player 1	1.5	0.543638	2.04364	2.04364
Player 2	1.5	0.734878	2.23488	2.23488
Player 3	1.5	1.019	2.519	2.519
Player 4	1.5	1.4752	2.9752	2.9752

Technology 3:

	Production(t=0)	Investment(t=0)	Production(t=1,up)	Production(t=1,down)
Player 1	1	0	1	0.930439
Player 2	1	0	1	0.915497
Player 3	1	0	1	0.888464
Player 4	1	0	1	0.849663

Technology 4:

	Production(t=0)	Investment(t=0)	Production(t=1,up)	Production(t=1,down)
Player 1	0	0.752366	0.752366	0.752366
Player 2	0	0.734878	0.734878	0.734878
Player 3	0	0.715204	0.715204	0.715204
Player 4	0	0.685119	0.685119	0.685119

Technology 5:

	Production(t=0)	Investment(t=0)	Production(t=1,up)	Production(t=1,down)
Player 1	0	1.06546	1.06546	1.06546
Player 2	0	1.0436	1.0436	1.0436
Player 3	0	1.019	1.019	1.019
Player 4	0	0.981399	0.981399	0.981399

Technology 6:

	Production(t=0)	Investment(t=0)	Production(t=1,up)	Production(t=1,down)
Player 1	1	1.06546	2.06546	2.06546
Player 2	1	1.0436	2.0436	2.0436
Player 3	1	1.019	2.019	2.019
Player 4	1	0.981399	1.9814	1.9814

Technology 7:

	Production(t=0)	Investment(t=0)	Production(t=1,up)	Production(t=1,down)
Player 1	1	0.623822	1.62382	1.55073
Player 2	1	0.590433	1.59043	1.52583
Player 3	1	0.569906	1.56991	1.48077
Player 4	1	0.534292	1.53429	1.41611

Technology 8:

	Production(t=0)	Investment(t=0)	Production(t=1,up)	Production(t=1,down)
Player 1	1	1.58728	2.58728	2.58728
Player 2	1	1.55813	2.55813	2.55813
Player 3	1	1.52534	2.52534	2.52534
Player 4	1	1.4752	2.4752	2.4752

Technology 9:

	Production(t=0)	Investment(t=0)	Production(t=1,up)	Production(t=1,down)
Player 1	0	1.602	1.602	1.50573
Player 2	0	1.5759	1.5759	1.48313
Player 3	0	1.56991	1.56991	1.48077
Player 4	0	1.53429	1.53429	1.41611

Technology 10:

	Production(t=0)	Investment(t=0)	Production(t=1,up)	Production(t=1,down)
Player 1	0	1.58728	1.58728	1.58728
Player 2	0	1.55813	1.55813	1.55813
Player 3	0	1.52534	1.52534	1.52534
Player 4	0	1.4752	1.4752	1.4752

The expected profit as well as total production/investment for each player is shown in Table 3.3.

Table 3.3: Equilibrium Price, Total Production, Investment and Expected Profit

	Price	Production(t=0)	Investment(t=0)	Production(t=1,up)	Production(t=1,down)	E(profit)
Player 1		7	8.827306	15.639216	14.863751	718.515985
Player 2		7	8.839549	15.839548	15.013172	728.857332
Player 3		7	8.9627	15.962704	15.283468	740.392629
Player 4		7	9.1763913	16.176388	15.671511	755.027756

P(t=0) = 62
P(t=1,up) = 31.3831
P(t=1,down) = 24.1681

The solution to this example underscores the tradeoff between investment level and production cost. For example, technology 1 may be considered similar to a "gas fired generator" with low capital costs but high production cost. The converse is true for technology 2, which is similar to a "hydro generator". Since the production cost associated with technology 2 is low, players produce as much as they can in either scenario. Consequently for either scenario, production due to technology 2 is limited by its capacity, and hence the output from technology 2 for either scenario is the same. This is akin to serving a "base load" of electricity demand. Beyond this demand, technology 1 is used to meet the remaining requirement. Finally, since cost functions are indexed in descending order, expected profits are increasing from first player to last player in this example.

2.3 A Two-Stage Game with Expected Scenarios (GES)

While there have been some previous attempts at modeling such stochastic equilibria (e.g., Haurie, *et al* [1990], Pakes and McGuire [1994]), numerical solution of such games is not as common as their deterministic counterparts. Given this tendency to model games in a deterministic setting, a natural question arises: What differences might one observe in the performance of the players if they all played as though the future were well represented using deterministic models in which expected values were used for decision making? That is, suppose that the players replaced the scenario based inverse demand curves with the expected inverse demand curves, then how do the firms fare in a stochastic environment? We will investigate the consequences of a modeling assumption that replaces random variables by their expectation, although the real world may involve uncertainty.

Such issues have already been addressed in the stochastic programming (SP) literature for the case of monopolies. There it is shown that by replacing the probabilistic scenarios by an expected scenario, the monopolist may paint too rosy a picture of future profits, when in fact, the expected profits obtained from such expected value models can be shown to be lower than that obtained from a stochastic decision model (see e.g., Birge and Louveaux [1997]).

We study this issue for oligopolies for the above mentioned game (GES). In the first step investment decisions are made from an expected scenario. In the second step, given these investment decisions, players choose their optimal production decisions. Formally it is defined in the following manner. Let i be a player from a set

$N = \{1,2,...,n\}$, n is finite. In *the first step* the players view the future in the form of an expected scenario ($\bar{s}$) and make investment decisions as follows.

$$\bar{I}_i \in \arg\max\left\{-F_i(I_i)+[f_i(I_i,q_i^*(K_i^1,\bar{s}),Q^*(K_i^1,\{K_j^{1*}\}_{j\neq i};\bar{s}),\bar{s}]\,|\,I_i \in A_i\right\} \tag{5}$$

where $\quad q_i^*(K_i^{1*},\bar{s}) \in \arg\max\{f_i(\bar{I}_i\ ,q_i,Q^*(K^{1*};\bar{s}),\bar{s})\,|\,q_i \in B_i(K_i^{1*})\}$,

$Q^*(K^{1*};\bar{s}) = \sum_{i=1}^{n}\sum_{k=1}^{m} q_{i,k}^*(K_i^{1*},\bar{s})$ and,

$f_i(.) = P(Q^*(K^{1*};\bar{s}),\bar{s})\sum_{k=1}^{m} q_{i,k} - c_i(q_i)$, $\quad K_i^{1*} = K_i^0 + \bar{I}_i$, and $\bar{s} = E(\tilde{s})$.

In *the second step* the players choose production quantities for period 0, and period 1. In fact, production decisions in period 0 are not affected investment decisions in period 0, whereas, production in period 1 are affected by investments in period 0. These two decision problems are as follows.

(Period 0): $q_i^*(K_i^0)$ for all $i \in N$ such that when $q_j \equiv q_j^*$, $j \neq i$ then

$$q_i^*(K_i^0) \in \arg\max\{P(Q^*(K^0))\sum_{k=1}^{m} q_{i,k} - c_i(q_i)\,|\,q_i \in B_i(K_i^0)\}, \tag{6}$$

where $Q^*(K^0) = \sum_{i=1}^{n}\sum_{k=1}^{m} q_{i,k}^*(K_i^0)$ and $B_i(.)$ is the production set for player i and K_i^0 is vector of available capacities from different technologies at time 0 for player i.

(Period 1): Production levels are $q_i^*(K_i^{1*},s)$ such that

$$q_i^*(K_i^{1*},s) \in \arg\max\{f_i(\bar{I}_i,q_i,Q^*(K^{1*};s),s)\,|\,q_i \in B_i(K_i^{1*})\}, \tag{7}$$

where $Q^*(K^{1*};s) = \sum_{i=1}^{n}\sum_{k=1}^{m} q_{i,s,k}^*(K_i^{1*},s)$, $\ f_i(.) = P(Q^*(K^{1*};s),s)\sum_{k=1}^{m} q_{i,s,k} - c_i(q_i)$

$K_i^{1*} = K_i^0 + \bar{I}_i$ and $\bar{I}_i$ is determined by the first step of this game.

We are interested in the performance of players under the two alternative games (GPS and GES). Note that even in the presence of uncertainty, the players may play the GES process. Nevertheless, their performance should still be evaluated in a probabilistic setting as explained above. This performance may be expected to be different from that obtained via the GPS process. The question we pose is: Is there a consistent bias? That is, do players playing one of the games always fare better than players involved in the other? Within the SP literature for the case of monopolies, there is a well known result that a model with probabilistic scenarios provides better performance than a model with expected scenarios. The difference between these values is referred to as the "value of the stochastic solution". However an analog of this result in the game setting is unknown.

As in Example 3.1, suppose that the demand scenarios are modeled by Figure 3.1. One may interpret δ in Figure 3.1 as a volatility level and our characterization suggests that the profitability of the players depends on the game they choose and the volatility of the market. The following lemma applies to both GPS and GES.

Lemma 1: Consider a market with n firms facing a future that is described by two states: an "up" state, and a "down" state. In this market, we assume that the inverse demand curve is linear in either state, and these curves have the same slope, regardless of the state. Moreover, suppose that cost function of each firm, denoted c_i, can be expressed as the sum of convex separable functions; that is, $c_i\left(\sum_k q_{i,k}\right) = \sum_k g_{i,k}(q_{i,k})$, where $g_{i,k}$ are convex and differentiable. Then under either the GPS game or the GES game, each player produces a greater quantity in the "up" state than in the "down" state.

Proof: Let $q_{i,k,u}$ and $q_{i,k,d}$ denote the quantities produced by firm *i* using technology *k* in the "up" and "down" states respectively. The claim is that $\sum_k q_{i,k,u} \geq \sum_k q_{i,k,d}$ for all *i*. Contrary to the claim, suppose that there exists an index *I* for which $\sum_k q_{I,k,u} < \sum_k q_{I,k,d}$. Then, there must exist a technology index *K* such that $q_{I,K,u} < q_{I,K,d}$. Now the KKT conditions for each firm imply that

$$P(Q_s) + q_{I,K,s} P'(Q_s) - g'_{I,K}(q_{I,K,s}) - \mu_{I,K,s} = 0, \qquad s = u, d$$

where $Q_s = \sum_{i,k} q_{i,k,s}$. Since $q_{I,K,u} < q_{I,K,d}$, it follows that $\mu_{I,K,u} = 0$. It follows that

$$P(Q_u) + q_{I,K,u} P'(Q_u) - g'_{I,K}(q_{I,K,u}) = 0 .$$

On the other hand, $\mu_{I,K,d} \geq 0$ implies that

$$P(Q_d) + q_{I,K,d} P'(Q_d) - g'_{I,K}(q_{I,K,d}) \geq 0 .$$

Using the KKT conditions for the pair *I,K*, the supposition that $q_{I,K,u} < q_{I,K,d}$, and $P'(Q_u) = P'(Q_d) < 0$, we have

$$P(Q_u) = -q_{I,K,u} P'(Q_u) + g'_{I,K}(q_{I,K,u}) < -q_{I,K,d} P'(Q_d) + g'_{I,K}(q_{I,K,d}) \leq P(Q_d) .$$

Due to the monotonicity of the (linear) inverse demand curve, we conclude that under our supposition, $Q_u > Q_d$. But this inequality implies the existence of another pair, say *J,T* such that $q_{J,T,d} < q_{J,T,u}$. Using the same arguments as above, we then come to the conclusion that $Q_d > Q_u$. These contradictory conclusions imply that there cannot exist any index I for which $\sum_k q_{I,k,u} < \sum_k q_{I,k,d}$, and hence the result. □

The following proposition also holds for both games.

Proposition 1: Assume that the probability of the "up" state is at least as high as the probability of the "down" state and consider a case in which all players experience identical quadratic costs and capacities. Moreover, assume that equilibrium involves unconstrained production in the low-demand state. Then players are better off as demand volatility increases.

Proof: This proof is for the case in which each firm has a single technology. The proof for the case of multiple technologies is similar. Given $P_0 = D - Q_0$, $P_u^1 = D + \delta - Q_u^1$, $P_d^1 = D - \delta - Q_d^1$, the Lagrangian function for player i is;

$$L_i = (D - \sum_i q_i^0)q_i^0 - c(q_i^0) - F(I_i^0) + u[(D + \delta - \sum_i q_{i,u})q_{i,u} - c(q_{i,u})] + d[(D - \delta - \sum_i q_{i,d})q_{i,d} - c(q_{i,d})]$$
$$+ \lambda_{i,0}(K_i^0 - q_i^0) + \lambda_{i,1}(K_i^0 + I_i^0 - q_{i,u}) + \lambda_{i,2}(K_i^0 + I_i^0 - q_{i,d}) \ .$$

The total derivative of above expression with respect to δ is;

$$\frac{dL_i}{d\delta} = \frac{\partial L_i}{\partial q_i^0}\frac{\partial q_i^0}{\partial \delta} + \frac{\partial L_i}{\partial q_{i,d}}\frac{\partial q_{i,d}}{\partial \delta} + \frac{\partial L_i}{\partial q_{i,u}}\frac{\partial q_{i,u}}{\partial \delta} + \frac{\partial L_i}{\partial I_i^0}\frac{\partial I_i^0}{\partial \delta} + \frac{\partial L_i}{\partial \delta} + \sum_{j \neq i} \frac{\partial L_i}{\partial q_{j,d}}\frac{\partial q_{j,d}}{\partial \delta} + \frac{\partial L_i}{\partial q_{j,u}}\frac{\partial q_{j,u}}{\partial \delta} + \frac{\partial L_i}{\partial q_j^0}\frac{\partial q_j^0}{\partial \delta}$$

$$= 0 + 0 + 0 + 0 + u.q_{i,u}[1 - \sum_{j \neq i} \partial q_{j,u} / \partial \delta] - d\, q_{i,d}[1 + \sum_{j \neq i} \partial q_{j,d} / \partial \delta]$$ (by first order necessary conditions),

<u>Case 1.</u> Up-state production is unconstrained.

Let $c'(q) = a + bq$, $F'(I) = g + hI$, with non-negative coefficients *a*, *b*, *g*, *h*. From the first order necessary conditions

$D - 2q_i^0 - a - bq_i^0 - \sum_{j \neq i} q_j^0 = 0$. Hence $q_i^{0^*} = (D-a)/(n+1+b)$. Similarly,

$q_{i,u}^* = (D + \delta - a)/(n+1+b)$, $q_{i,d}^* = (D - \delta - a)/(n+1+b)$. Thus $\partial q_{i,u}^* / \partial \delta = 1/(n+1+b)$,

$\partial q_{i,d}^* / \partial \delta = -1/(n+1+b)$, and by the result of Lemma 1 and $u \geq d$, it is easy to see that

$d\Pi_i / d\delta = dL_i / d\delta > 0$.

<u>Case 2.</u> Up-state production is constrained.

Then $q_{i,u}^* = K_i^0 + I_i^{0^*}$, and $F'^*(.) = g + h(q_{i,u}^* - K_i^0)$, and the first order condition becomes

$u[D + \delta - 2q_{i,u} - a - bq_{i,u} - \sum_{j \neq i} q_{j,u}] - g - h(q_{i,u} - K_i^0) = 0$, which implies

$q_{i,u}^* = [u(D + \delta - a) - g + hK_i^0]/[u(n+1+b) + h]$. Thus $\partial q_{i,u}^* / \partial \delta = 1/(n+1+b+h/u)$ and

$dL_i / d\delta = uq_{i,u}^*[1 - 1/(n+1+b+h/u)] - dq_{i,d}^*[1 - 1/(n+1+b)] > 0$. □

The following example illustrates a comparison of GPS and GES.

Example 2: The aim of this example is to compare these two games in terms of their performance as volatility changes. We continue with a setting similar to Example 1, but this time only two technologies are available for each player. Specifically, initially each player has 3 units of the technology 1 capacity, (i.e. $K_{i,1} = 3$), and has 2 units of technology 2 capacity, (i.e. $K_{i,2} = 2$). Investment and production cost functions are $F_{i,1}(.) = C_{i,1} I_{i,1}$ *and* $c_{i,1}(.) = C_{i,1} q_{i,1}^2$, and $F_{i,2}(.) = C_{i,2} I_{i,2}^2$ *and* $c_{i,2}(.) = C_{i,2} q_{i,2}$, $C_{i,m} = i + 1$, for technologies 1 and 2 respectively, and i =1,2,3,4,5. We investigate cases when δ assumes the following values: 0,5,10,15,20,25,30. In GES, in the *first step* each player solves (5) with $\bar{s} = E(\tilde{s})$, in that case s=1. From here each player implements this investment decision to the following *second step* optimization problem.

$$
\begin{array}{lll}
\max\ \Pi_i(q^t_{i,s}, \bar{I}^t_i, K^t_{i,s}) & & (4.1')\\
\textit{subject to} & & \\
q^t_{i,s,k} - K^t_{i,s,k} \le 0 & t=0,1,\ \forall\ i,s,k & (4.2')\\
Q^t_s - \sum_{i,k} q^t_{i,s,k} = 0 & t=0,1\ ,\ \forall i,s,k & (4.3')\\
q^t_{i,s,k} \ge 0\ ,\ K^t_{i,s,k} \ge 0 & t=0,1\ ,\ \forall i,s,k & (4.4')\\
\bar{I}^t_{i,s,k} \ge 0 & t=0\ ,\quad \forall i,s,k & (4.5')\\
K^{t+1}_{i,s,k} - K^t_{i,s,k} - \bar{I}^t_{i,s,k} = 0 & t=0\ ,\quad \forall i,s,k & (4.6')
\end{array}
$$

In Figure 3.2, we illustrate player 3's expected profit values with respect to volatility changes under both games. Because all players have similar cost functions, we observe similar figures for other players too. Note that for each volatility level, GES dominates GPS. In GPS, there is an investment opportunity as demand volatility increases. In contrast, greater demand volatility has no impact on investment decisions in GES since the expected level of demand is unchanged. As demand volatility increases, players in the GPS game increase their investments in order to take advantage of higher demand in the good demand state; they can leave extra capacity unutilized in the low demand state.

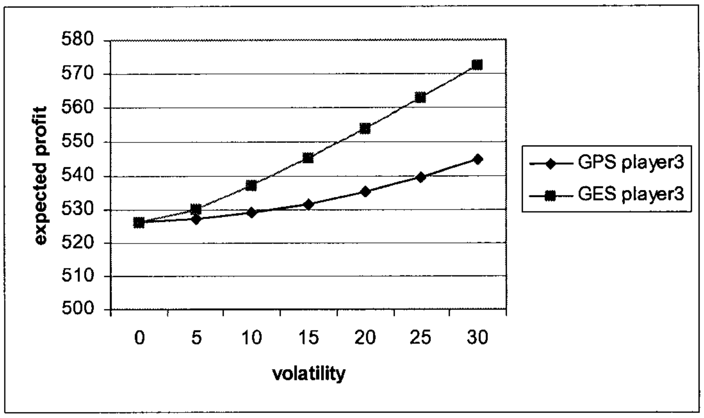


Figure 3.2: Player 3's performance under these games

2.4 A two-stage Hybrid Game (HG)

In this subsection, we show that despite the possible advantage of GES over GPS, the former is unstable because it creates a situation in which one of the players may have an incentive to not play by the rules of GES. In order to see this, we consider a hybrid game (HG) which is a combination of GPS and GES in the following manner.

In the first step of the hybrid game all players assume that the others play according to the first step of GES. However, player i chooses to deviate by planning for investments by solving a decision model in which the probabilistic scenarios are included. Once the investments are in place, the second step of the game proceeds as before. It is not difficult to see that as long as GPS and GES have solutions, the hybrid game also has solutions. In such a hybrid game, we can show that player i will always be better off.

Proposition 2 : Suppose that a player who deviates from GES and plays GPS and others play GES, this hybrid is named as HG, then this player's expected profit function is ordered as $\Pi_{GES} \leq \Pi_{HG}$ under the assumptions made above regarding demand and technology conditions.

We will give a sketch of the proof. Since all non-deviating players have committed their decisions then the game reduces to a single player stochastic program. Applying a basic result (proposition 1 of Birge and Louveaux [1997] on page 140) from stochastic programming it follows that the expected profit of the deviating player in the stochastic program exceeds that from the expected value model. The above result and the corresponding arguments hold in the multi-stage game (see section 3).

Example 3: We assume that there are four players having two types of technologies. Players are supposed to make investment and production decisions in a two-stage model. Also we assume same cost functions and demand scenarios as defined in Example 2. Now let one player, say player 1, play HG, and other players play GES. In this case we get following figure for player 1's performance under all above defined games. It is clear that by deviating player 1 improves his performance as suggested by Proposition 2. Thus, even though it may be more profitable to play within a GES game, it creates an unstable situation, and firms may therefore be expected to play the GPS game.

With the insight gained from the HG, let us reconsider the conclusions from Example 2. If only one player reacted to the positive volatility, while the other players retained their GES investment levels, then this player would increase their investment and their expected profit would rise. This increase in expected profit corresponds to the value of the stochastic solution, referred to earlier. But in a multi-player game setting, players' payoffs are interdependent. In fact all of the players will increase their investments and their outputs. The net effect is that prices and expected profits are lower in the GPS game than in the GES game. This can be viewed as a prisoners' dilemma situation. Players in the GPS game end up being worse off by pursuing what is individually optimal (given strategies of rival players) compared to players in the GES game who behave sub-optimally from an individual point of view, but more nearly optimally from the point of view of collective profits.

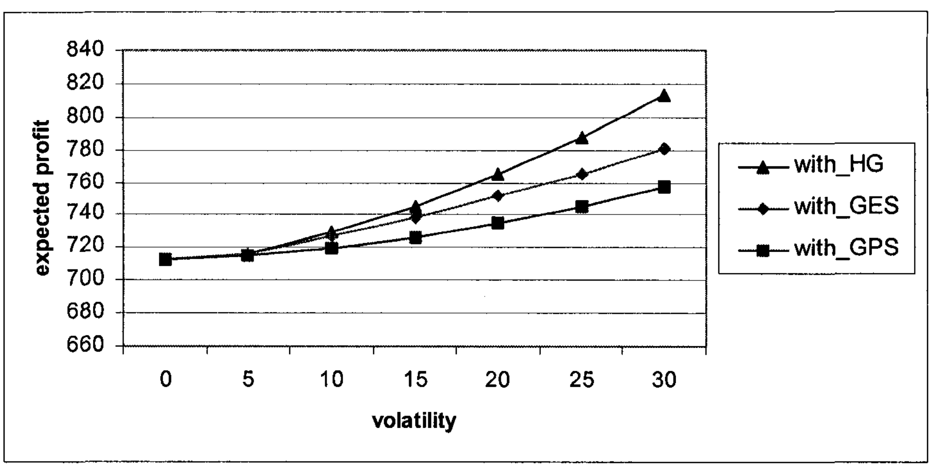


Figure 3.3: Player 1's performance under all games

3. Multi-stage Stochastic Programming Games

In this section we study multi-stage games that are generalizations of the two stage games presented in section 2. However in the interest of analytical tractability we impose certain conditions on the behavior of the players. Following this discussion we will provide the games and some associated results.

3.1 A Model of Demand and Supply in the Multi-stage Case

In multi-stage formulations the demand function may be modeled as a discrete-time random walk process. The price-quantity relationship is assumed to be governed by the form $aP_t + bQ_t = D + \tilde{\delta}_t$, and for t=0, $\tilde{\delta}_0 = 0$ and for t=1,2,3,…,T , $\tilde{\delta}_t = \tilde{\delta}_{t-1} \pm \delta$, where $a, b, D > 0$ are given constants. $P, Q \in \Re_+$ and denote the market price and total output, respectively, and $\tilde{\delta}$ is the random variable. We assume that the outcomes of the random variable take two values: $+\delta$ for the up-state, $-\delta$ for the down-state with some probability, and $\delta \geq 0$. Figure 3.4 illustrates the behavior of a demand function modeled as a random walk.

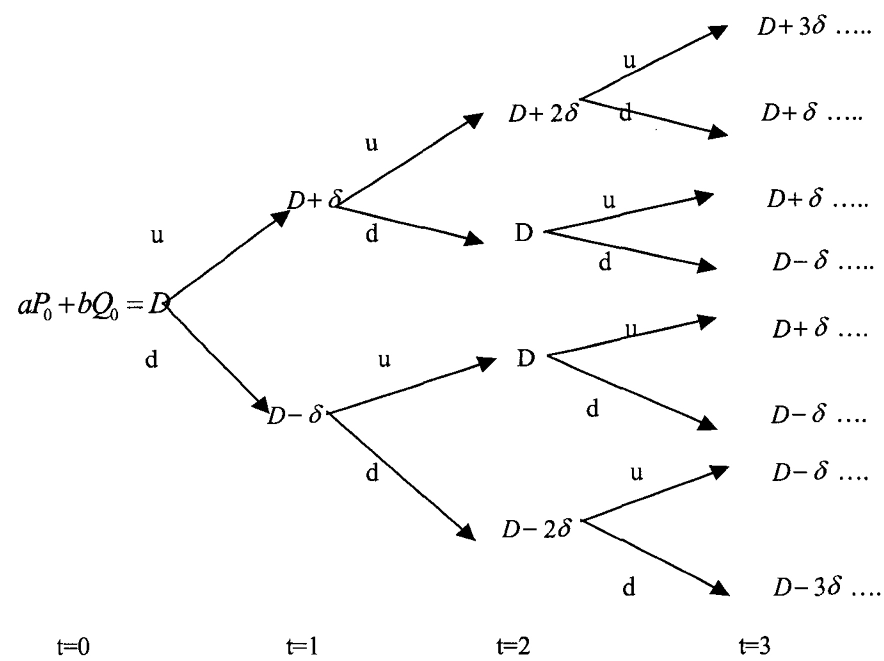


Figure 3.4: Discrete-time random walk representation of demand function

Conditional probabilities are given on the arcs, with the upside probability denoted as '*u*', and the downside probability denoted as '*d*' and of course *u*+*d*=1. The probability distribution of $\tilde{\delta}_t$ depends only on the state $\delta_{t-1}+D$, and in each period up-state and down-state probabilities are independent of what happened in the previous periods. Hence $\tilde{\delta}_t$ follows a Markov process with independent increments. It is easy to show that continuous time representation of this process is the Brownian motion with drift (see Pyndick and Dixit [1997]).

The scenario tree used in section 2 is highly simplified, because there was only one future stage with two possible outcomes. In the multi-stage formulation, the scenario tree can become considerably more complex. In figure 3.4, we give an example of scenario tree, which captures the sequence in which information is revealed. Also, in the class of models we are considering (i.e., Stochastic Programming) decisions are not allowed to alter the underlying stochastic process. So the underlying stochastic process is assumed to be exogenous to the decision process.

The supply conditions for a firm are characterized by its production and investment cost functions. We allow firms to have multiple technologies which have different characteristics. For example, in one technology production cost may be higher than investment cost, for some technologies it may be vice-versa, even some technologies may have negligible (almost zero) production cost with fixed investment cost. Also we assume that these cost functions are convex, quadratic and increasing. These strictly convex investment cost functions may stem from adjustment costs, which are due to costs of installing and/or removing equipment (see Mussa [1977] and Reynolds [1986]).

3.2 A Multi-Stage Game with Probabilistic Scenarios

Multi-stage games are multi-period extension of two-stage games. At any given period t, t is finite, players simultaneously choose investment and production decisions from multiple technologies. As in the previous section, investment in one period will only become available as productive capacity in the next period. There are more general lags that can also be modeled within this framework. However we omit such generalizations in the interest of brevity.

The setting for our formulation is one in which each player can only observe production output of the other players through the market price. We formulate this game recursively and non-recursively. The recursive formulation presented is based on the formulation of Haurie, *et al* [1990], and is included in the Appendix for the sake of completeness.

On the other hand the non-recursive formulation is as follows. Let $M(o)$ represent scenarios passing through node o of the scenario tree. As before let $I_{i,s}^t, q_{i,s}^t \in \Re_+^m$ denote investment and production quantities associated with m technologies. The total expected profit for a player is denoted $\Pi_i = \sum_s p_s \Pi_{i,s}$, where p_s is the probability of each scenario. Formally we define this game as follows.

For a given scenario $s \in S = \{1,2,3,\ldots,\omega\}$, each player i solves following problem:

$$\underset{q_{i,s}^t, I_{i,s}^t}{Max} \sum_t \Pi_{i,s}^t(q_{i,s}^t, I_{i,s}^t, K_{i,s}^t, P_s^t) \quad , \; t = 0,1,2,3,\ldots, \mathrm{T}\,,$$

and subject to $K_{i,s}^{t+1} = K_{i,s}^t + I_{i,s}^t$, (8)

and subject to $E(I_{i,s}^t \mid M(o)) = I_i^t$, $q_{i,s}^t \le K_{i,s}^t$, where in (8) we have used a standard one period lag. However if there are more complicated time lags, then (8) need to be modified to suit the situation being modeled; nevertheless the remainder of the model as well as associated algorithms remains unchanged.

Other applications in which this formulation may be advantageous include situations in which a player's objective is not separable by time period. For instance if the players' objectives include risk preferences then such models may not be separable, and once again a non-recursive formulation becomes necessary.

A solution algorithm for the non-recursive formulation is as follows:

i) The Karush- Kuhn-Tucker conditions for the problem are as follows: $\forall i,t,s$

$$L_{i,s} = \sum_{t}^{T} \beta(t)\{P_s^t(Q_s^t)q_{i,s}^t - c_i(q_{i,s}^t) - F_i(I_{i,s}^t) + \alpha_{i,s}^t[K_{i,s}^t - q_{i,s}^t] + \lambda_{i,s}^t[E(I_{i,s}^t) - I_i^t]\}$$

$$q_{i,s}^t \frac{\partial L_{i,s}}{\partial q_{i,s}^t} = 0 \ , \ I_{i,s}^t \frac{\partial L_{i,s}}{\partial I_{i,s}^t} = 0 \ , \ \alpha_{i,s}^t \frac{\partial L_{i,s}}{\partial \alpha_{i,s}^t} = 0$$

$$q_{i,s}^t \le K_{i,s}^t \ , \quad I_{i,s}^t - E(I_i^t) = 0$$

$$K_{i,s}^t = K_{i,s}^{t-1} + I_{i,s}^{t-1} \ , \quad Q_s^t - \sum q_{i,s}^t = 0$$

$$K_i^0 = \kappa \ge 0 \ , \ q_{i,s}^t \ge 0 \ , \ I_{i,s}^t \ge 0 \ ,$$

where $(\alpha_{i,s}^t \ge 0)$, $(\lambda_{i,s}^t \ge 0)$ denote for dual prices of inequalities and non-anticipative conditions respectively, and $\beta(t)$ is a discount factor.[4]

ii) Use the PATH solver for a complementarity problem.

Proposition 3: For the case of a one-period lag between investment and production, the recursive formulation is equivalent to the non-recursive formulation for oligopolistic games.

Proof: The recursive formulation for the multi-period GPS game (hereafter NF) for a firm i may also be defined in the following manner:

[4] In all examples we assume the discount factor to be 1, since the problems involve a fairly small number of stages.

$$\max \sum_{o} p_o \Pi_o^i(x_o^i)$$

$$s.t. \sum_{m \in A_o} I_{o,m}^i + K_0^i \geq q_o^i \quad , \quad \forall o \tag{9}$$

$$q_o^i, x_o^i \geq 0 \ , \ K_0^i > 0$$

where, $x_o^i = (I_o^i, q_o^i)$, $\Pi_o^i(.) = \mathrm{P}_o q_o^i - c(q_o^i) - F(I_o^i)$, and A_o denotes for ancestors of node o.

The non-recursive formulation of the multi-stage GPS game (hereafter SF) for a firm i may also be defined as follows:

$$\max \sum_{s} p_s \sum_{t} \Pi_{ts}^i(x_{ts}^i)$$

$$s.t. \sum_{m \in A_{st}} I_{o,m}^i + K_0^i \geq q_{ts}^i \quad , \quad \forall t,s$$

$$x_{ts}^i = E(\tilde{x}_t^i) \tag{10}$$

$$q_{ts}^i, x_{ts}^i \geq 0 \ , \ K_0^i > 0$$

where, $x_{ts}^i = (I_{ts}^i, q_{ts}^i)$, $\Pi_{ts}^i(.) = \mathrm{P}_{ts} q_{ts}^i - c(q_{ts}^i) - F(I_{ts}^i)$

Let x_o^i be a solution for NF and let M_o be the scenarios that pass through the node o, then for $t_o = t$ define the following relation; $x_o^i = x_{t_o s}^i$ for all $s \in M_o$. That means $I_{t_o s}^i = I_o^i$ and $q_{t_o s}^i = q_o^i$ by definition of x_o^i. But these two equalities are just non-anticipativity constraints of SF, and so that $\sum_t \Pi_{ts}^i(.) = \Pi_o^i(.)$, then we see that proposed solution is a feasible solution of SF. By existence and uniqueness theorems this feasible solution is same as the solution of NF. The reverse is similar. Because, if we have non-anticipative solutions, then the fact that the I, q are non-anticipative imply that they are constant at any node, thus letting us set the value of the node variable for I, q to assume the same value as that of all scenarios passing through node o.□

Proposition 4: The non-cooperative games defined in this paper together with the above demand and technology conditions admit a unique solution.

For a proof see Haurie, *et al* [1990] or see the proofs of theorems 7.1 and 7.7 in Friedman [1977].

***Example 4*:** This example extends two stage games to multi-stage games with multiple players having multiple technologies. Specifically, the market is comprised of four players having two different technologies. Players make optimal investment and production decisions along five stages. The data in Table 3.4 shows each player's initial capacities and investment and production cost coefficients from technologies 1 and 2 respectively.

Table 3.4: Cost coefficients and initial capacities from each technology

	Invest./Product. Cost Coeff.	Initial Capacity
Player 1	(6,6)	(1.5,1.5)
Player 2	(6,5)	(1.5,1.5)
Player 3	(5,4)	(1.5,1.5)
Player 4	(5,3)	(1.5,1.5)

For technology 1 (m=1), investment and cost functions are in the following form: $F_{i,m}(.) = C_{i,m}I_{i,m}$ *and* $c_{i,m}(.) = C_{i,m}q_{i,m}^2$. This technology may represent "gas-fired" generator, in which production cost is significantly higher than the investment cost. Player 1 and player 2 have the same cost coefficient which is slightly higher than the cost coefficients of player 3 and player 4 who have equally costly technology. The cost functions from technology 2 (m=2) are defined as $F_{i,m}(.) = C_{i,m}I_{i,m}^2$ *and* $c_{i,m}(.) = C_{i,m}q_{i,m}$.

This technology may be considered as "hydro" generator since investment cost is much more than production cost. Each player i (=1,2,3,4) predicts a market with discrete time random walk scenarios as depicted in Figure 3.4, where D=90 , δ=5, u=0.5 is the probability of the "up" state, and d=0.5 is the probability of "down" state.

We model this as a game with probabilistic scenarios and compute the unique S-adapted open-loop equilibrium. As shown in Figure 3.4, the equilibrium prices mimic the scenario tree structure. The expected price and standard deviation of prices for each period of this game are provided in Table 3.5. In Figure 3.6 we provide a histogram of prices in the final period, based on an interval width equal to two.

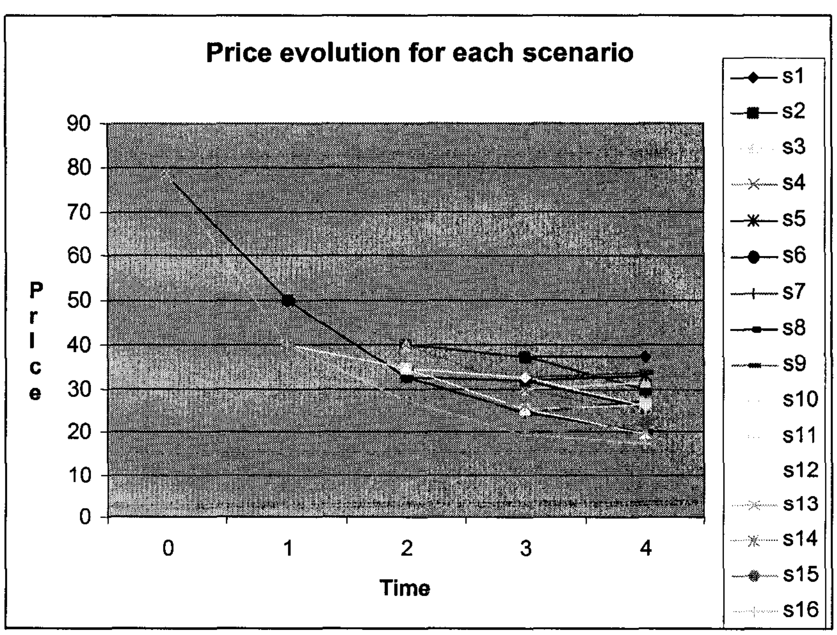


Figure 3.5: Price trajectories for each scenario

Table 3.5: First and second moments of price series for each period

time	Expected Price	Stand.Deviation
0	78	-
1	45.0533	7.071
2	33.5929	5.3717
3	28.3655	5.5779
4	26.1489	5.6686

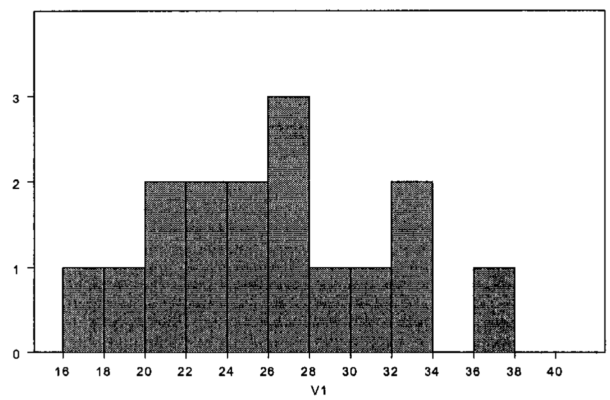


Figure 3.6: Histogram of prices for t=4

A potential alternative to computation of the S-adapted open-loop equilibrium via stochastic programming is to compute a feedback equilibrium using Dynamic Programming. A natural state vector for such an exercise would be the demand state and the vector of capacities for the 2 technologies for each of the four players. This yields a 9-dimensional state vector. If this game could be expressed as a finite time-horizon, linear-quadratic (LQ) game then well-known recursive algorithms for computation of linear feedback equilibrium decision rules could be applied (e.g., see Kydland [1977]). However, a LQ formulation does not permit constraints on decisions (e.g., production bounded below by zero and above by capacity) nor does it provide a good approximation of capacity constrained costs for Example 4. A quadratic payoff would imply that marginal production cost is linear in output for all output levels. This is a bad approximation of an increasing marginal cost curve that becomes vertical at output equal to capacity.

Another approach would be to discretize the state space by restricting capacity levels to belong to a finite set of m numbers. The DP approach would involve computation of a feedback equilibrium value function for each player for each possible state. The number of possible states in period t for Example 4 is $2^t m^8$ (2^t is the number of demand states in t, m^8 is the number capacity vectors for 2 technologies and 4 players). For example, if m=10 then there are 1.6 billion possible states in the final period (t=4), 800 million possible states in the penultimate period (t=3), and so on.[5] This illustrates the curse of dimensionality for a DP approach. Application of the DP approach to this seemingly modest-sized problem turns out to involve an enormous computational problem.

[5] Of course, if m is smaller the dimensionality is reduced. But a courser grid of capacity values may be an inferior approximation of the true state space. A discrete state space may also introduce new problems of multiple equilibria. Berry and Pakes [1993] utilize a probabilistic investment formulation that smoothes out the discreteness and avoids the multiplicity problem. However, this formulation does not eliminate the "curse of dimensionality".

4. Conclusions

In this paper we have studied several alternative games under uncertainty. These include Games with Probabilistic Scenarios (GPS), Games with Expected Scenarios (GES), and a Hybrid Game (HG). The HG formulation provides the argument against GES, even though there are instances in which all players may be better off by choosing to play the GES over GPS. Thus under uncertainty, it is plausible to consider the GPS formulation. We also show that if all players are identical, and the probability of a higher inverse-demand curve either equals or exceeds that of a lower inverse-demand curve, then players can earn greater expected profits as volatility increases. This result suggests that even in an increasingly volatile market, players may have an incentive to participate in the market. Given the surge in uncertainty within certain markets (e.g., energy), this result appears to be reassuring.

Our approach, based on stochastic programming, provides a computationally realistic approach to oligopolistic games under uncertainty. The key to this presentation is the so-called scenario formulation which allows us to state the games through a finite number of scenarios, coordinated by the non-anticipativity constraints. The same type of formulation holds for both two-stage, as well as multi-stage games, both of which are solvable using standard algorithms for linear complementarity problems (e.g. the PATH solver available on the NEOS server). The examples presented in this paper illustrate that this approach can address dynamic games that are clearly out-of-reach of dynamic programming, the common approach in the literature on dynamic games. The theory associated with infinitely-many scenarios within the stochastic programming setting is currently under investigation, and some initial results are available in Higle and Sen [forthcoming] and Casey and Sen [2002].

APPENDIX A

Recursive Formulation

Let the scenario set be $S=\{S^1,S^2,...,S^T\}$, and formed by all possible scenarios in a decision tree, for each time period, where $s^t \in S^t$ is a particular scenario at time *t*, s_-^t is the predecessor node of s^{t+1}, so that $s^{t+1} \mid s_-^t$ is the path at time *t*+1 given we were on s_-^t. In an extensive form each player's optimization problem is the one we defined at (4) with $t = 0,1,2\ldots,$T and the payoff function

$$\Pi_i = q_i^0.P(Q^0) - c_i(q_i^0) - F_i(I_i^0) + \beta^1 \sum_{\forall s^1 \in S^1} \theta(s^1)\{q_{i,s}^1.P(Q_s^1) - c_i(q_{i,s}^1) - F_i(I_{i,s}^1)\}$$
$$+\beta^2 \sum_{\forall s^2 | s_-^1 \in S^2} \theta(s^2 \mid s_-^1)\{q_{i,s}^2.P(Q_s^2) - c_i(q_{i,s}^2) - F_i(I_{i,s}^2)\} + \ldots + \beta^t \left\{ \sum_{\forall s^T | s_-^{T-1} \in S^T} \theta(s^T \mid s_-^{T-1})\{q_{i,s}^T.P(Q_s^T) - c_i(q_{i,s}^T)\} + G_i \right\}$$

Here $q_{i,s}^t, I_{i,}^t \in \Re_+^m$, β^t is the discount factor at time *t*, and $\beta^t = \prod_{l=1}^{t} \beta^l$, also $\theta(s^t \mid s_-^{t-1})$ is the conditional probability on the arc and G_i is the salvage value of the investments. One could let the discount factors be path dependent (on the scenarios) but that would detract from our current focus. We keep simplicity by maintaining fixed discount factors over alternative paths.

Then a recursive formulation of a multi-stage game with probabilistic scenarios (NF) for a player *i* defined as follows.

$\underset{q_i^0, I_i^0}{Max}\ \Pi_i^0(q_i^0, I_i^0, K_i^0, P^0) + E[f_i^1(q_i^1, I_i^1, K_i^1, P^1, \tilde{s}^1)]$, for t =1,2,3,…, T-1 , the value functions are $f_i^t(.) = \underset{q_i^t, I_i^t}{Max}\ \Pi_i^t(q_i^t, I_i^t, K_i^t, P^t, s^t) + E[f_i^{t+1}(q_i^{t+1}, I_i^t, K_i^{t+1}, P^{t+1}, \tilde{s}^{t+1} \mid s_-^t)]$, and

$f_i^T(.) = \Pi_i^T(q_i^T, K_i^T, P^T, s^T \mid s_-^{T-1}) + G_i$, $K_i^{t+1}(s^{t+1} \mid s^t) = K_i^t(s^t) + I_i^t(s^t)$, $I_i^t, q_i^t \in \Re_+^m$, m is the number of technologies.

CHAPTER 5

SUPPLY FUNCTION EQUILIBRIA WITH PIVOTAL SUPPLIERS

1. Introduction

The supply function equilibrium (SFE) concept has become a widely used tool to study the exercise of market power by sellers in multi-unit auction environments. SFE models assume that each seller submits a continuous supply schedule for divisible output to the auctioneer. Klemperer and Meyer (1989) (hereafter KM) characterize SFE in environments for which product demand is uncertain. They show that a SFE is the solution to a system of differential equations. For a n-firm symmetric model they show that there are multiple equilibria when the range of demand variation is bounded. Roughly speaking, these equilibria are contained in a range of prices between the Cournot price and the competitive price.

The SFE has found its widest application in the analysis of wholesale electricity auctions. Many of these auctions are run as uniform price, multi-unit auctions in which power sellers submit offer schedules indicating their willingness to supply. Offer schedules are often submitted a day ahead of the actual auction. Many papers have utilized the SFE concept to analyze various aspects of electricity auctions. Examples include Green and Newbery (1992), Newbery (1998), Rudkevich, et al (1998), Green

(1999), and Baldick and Hogan (2002). These papers consider a variety of extensions and modifications of the KM model, including production capacity constraints, asymmetric firms, potential entry, multi-step cost functions, and forward contracting.

A number of recent assessments of wholesale electricity market performance have emphasized the role of the extent of excess production capacity in the market and the ability of a single supplier to influence the market price by withholding production (see Joskow and Kahn (2001), Lave and Perekhodtsev (2001), Rothkopf (2001), Borenstein, Bushnell and Wolak (2002), and Perekhodtsev, et al (2002). The term "pivotal supplier" has been used to describe an electricity supplier that is able to dictate the price in the auction by withholding some portion of its production from the auction. One or more pivotal suppliers are most likely to be present in an auction when demand (or, load) is near its peak, when available production capacity in the market is limited relative to the peak load, and when suppliers' capacities are asymmetrically distributed.

As noted above, the SFE concept has been widely applied in the analysis of electricity markets. A number of studies have examined how production capacity constraints influence the range of equilibrium prices (e.g., see Green and Newbery (1992) and Baldick and Hogan (2002)). Yet these studies have not examined the potential role of the extent of excess capacity in the market on equilibrium prices, nor have they shown how the presence of pivotal suppliers affects predicted equilibrium supply functions and prices.

In this paper we formulate a simple model of a wholesale electricity auction for which the notion of a pivotal supplier has a natural interpretation. We assume that demand varies over time (during the trading period), and is inelastic. In the symmetric models, we consider the case in which players' marginal cost is fixed up to capacity. In another case we assume that suppliers have step marginal costs, and total capacity is equally divided among them. In the asymmetric model, we assume firms are different in capacities, and have a common marginal cost for production up to capacity. The market price is bounded by a price cap. By withholding output, a pivotal supplier can move the market price to the maximum price, or price cap for the market. As in other SFE models with bounded demand variation, there is a continuum of equilibria. We examine the connection between pivotal suppliers and the set of SFE. In symmetric and asymmetric versions of the model we show that when pivotal suppliers are present, the set of SFE is reduced relative to when no suppliers are pivotal. We show that the size of the equilibrium set depends on observable market characteristics such as the amount of industry excess capacity, the load factor, the number of suppliers, and the amount of low-cost base load production capacity. For example, as the amount of industry excess capacity falls and/or the load factor rises and/or the number of suppliers decreases and/or the low-cost base load capacity falls in which the base load is less than the off-peak load level, the set of SFE becomes smaller; the SFE that are eliminated are the lowest-priced, most competitive equilibria. The firm with the larger share of capacity has an incentive to deviate from a wider range of SFE, and it is the larger firm's deviation incentives that determine which SFE are ruled out as equilibrium. We also introduce the idea of "optimal deviation" from proposed (candidate) SFE and see how it fares with respect to the "simple deviation".

Section 2 of this paper gives the summary of pivotal electricity suppliers and supply function equilibrium models in the literature. Section 3 describes the model of supply function equilibrium used in this paper. Section 4 considers symmetric firms model. In section 5, this model is used to explain optimal deviations. Section 6 introduces asymmetric firms model. Section 7 expands the section 4 by introducing step marginal costs into the model. Section 8 concludes.

2. Background

2.1. Pivotal Electricity Suppliers

In this section we review the literature on pivotal electricity suppliers, and explain how their presence may create market power. The discussion is organized around measures that have been utilized to assess the impact of pivotal suppliers. These measures include the Hirfindahl-Hirschman Index, Pivotal Supplier Index, Supply Margin Assessment, and Residual Supplier Index.

The Hirfindahl-Hirschman Index (HHI) measures market power based on generators' market shares; see Cardell, et al (1997) for an application to the electricity industry. Formally it is calculated as the summation of squares of firms' market shares. Oligopoly models such as the Cournot model of output choice predict that a market with a higher HHI will have higher markups of price over marginal cost, reflecting greater market power. There is empirical evidence from cross sectional manufacturing data that supports this prediction. However, the HHI is unlikely to capture the specific load and capacity conditions that yield pivotal suppliers in electricity markets.

The Pivotal Supplier Index (PSI) (see Bushnell, et al (1999)) measures market power based on generators' pivotal status. If the residual demand (total market demand minus summation of total maximum capacity of other generators and total exported power) for a generating firm is positive then this firm can be considered as a pivotal supplier which can exercise market power. At a point in time, the PSI is a binary variable for a generator such that, if residual demand is greater than 0, PSI takes 1 and the generator is assumed to be pivotal, otherwise it becomes 0 and the generator is non-pivotal. The PSI for a generator is obtained by averaging its PSI's over time. Essentially the PSI measures the frequency of monopoly power due to pivotal status held by a given firm. However, a possible shortcoming of this measure is that it only captures the most extreme level of market power. Using this approach Bushnell, et al (1999) conclude that suppliers in the Wisconsin/Upper Michigan (WUMS) region exercised market power by raising the prices well above the competitive level. They noted that the sources of this market power stem from high concentration of capacity ownership within WUMS, the limited transmission capacity available for imports into WUMS, the relatively tight reserve margins within WUMS region, and other well known factors such as inelasticity of demand, and non-storability of the electricity. They also examine whether asset divestiture and transmission expansion increase the competitiveness of the WUMS market, and find that these remedies are not sufficient for that purpose because of the market structure of the WUMS.

The Supply Margin Assessment (SMA) was designed by the Federal Energy Regulatory Commission (FERC). This is a form of PSI applied to annual peak condition. During the peak hours, if a supplier is pivotal, then this supplier fails the SMA screen test. Blackburn (2001) illustrates an operation of SMA during a peak hour with the following example:

Total control area generation:	15000
Total transmission capacity (TTC):	3000
Total market capacity:	18000
Control area load:	16000
Supply margin:	2000

(that is, fixed system capacity minus the demand=18000-16000)

A supplier's installed generation:	2500
A supplier's uncommitted generation:	0

According to FERC's 20% benchmark rule (if the generator's market share is more than 20% then generator is automatically assumed to have market power) this supplier has no market power, because its market share is 16.7% (2500 divided by 15000), which is less than the benchmark. However this generator fails under the SMA test; by withholding 2000 units of its capacity of 2500, the firm can create a shortage. The main weakness of the SMA measure is that it restricts and disqualifies the supplier for a single peak hour and assumes perfect information about load and supply. It also does not account for operating reserve margin requirement. Currently the FERC does not apply this SMA test to any generator at the ISO New England, New York ISO, PJM and the California ISO, because these generators got approved market monitoring and mitigation.

A market concentration index formulation for electricity markets should take into account factors, which affect market outcomes. These factors may be demand, total available supply, and the capacity of large supplier(s). PSI and SMA definitely consider these factors, and show whether residual supply meets the market demand or not. However these measures do not tell us anything about the relative sizes of demand and residual supply.

The Residual Supply Index (RSI) is a generalized form of PSI. RSI was devised by the California Independent System Operator (CAISO) (see Sheffrin (2001, 2002)). It is calculated as the ratio of residual supply (total supply minus largest seller's supply) to the total demand. Here, total supply is the summation of total in-state supply capacity and total net import. Total demand is the sum of metered load and purchased ancillary service. The largest seller's supply refers to the difference of its capacity and its contract obligation to load. If residual supply (total capacity less capacity of largest supplier) is sufficient to meet the market demand, then it is likely to reach competitive market outcomes. Using summer 2000 peak hourly data from the California PX market, Sheffrin (2001) shows that there is significant negative correlation between the Lerner Index ((price - marginal cost)/price) and RSI. The lower RSI is, the higher the price-cost markup. She finds that when RSI is about 1.2, the average price-cost markup is about 0, which is the competitive market benchmark.

Perekhodtsev, et al (2002) formulate and analyze a game theoretic model in which symmetric, capacity constrained firms submit offers to supply into a uniform price auction. They assume a fixed, inelastic demand. Their aim is to assess the role that pivotal suppliers play in price formation. They restrict attention to simple bidding strategies in which a firm bids either a "Low" price equal to marginal cost or a "High" price equal to the price cap. Equilibrium bidding involves mixed strategies in which each firm bids either low or high with specific probabilities. The equilibrium probability that the price is high depends on the supply margin, the difference between industry capacity and the fixed demand (load). As the supply margin increases (excess capacity decreases) the expected price in equilibrium falls. The presence of a single pivotal supplier is associated with a high price in their model. They also discuss the notion of a pivotal group of firms – a group of firms whose total capacity exceeds the supply margin. They show that market power gradually declines as the number of firms that are jointly pivotal rises. To examine the role of pivotal suppliers, they assess how observed price-cost margins in the California wholesale electricity market during late 2000 vary with the number of pivotal suppliers in the market. They find that price cost margins were higher the fewer the number of pivotal suppliers, as their model predicts[6].

2.2. Supply Function Equilibrium Models

A supply function, specifying the quantity that a firm is willing to produce as a function of price, may be viewed as a firm's strategy in a game. A model utilizing

[6] It should be noted that their theoretical conclusions are based on a very simple model with only two

strategies of this type was first formulated by Grossman (1981), and later studied by Hart (1985). They considered Nash equilibria in supply function strategies, or SFE, in the absence of uncertainty about demand.

Klemperer and Meyer (1989) extended the supply function model to allow uncertainty about demand. Firms are uncertain about the level of demand when they submit their supply functions. A uniform market price is determined by the intersection of the realization of the demand function and the aggregate supply function. KM note that a supply function strategy affords a firm greater flexibility, and correspondingly greater profits, than a fixed price or a fixed quantity strategy when demand is uncertain. KM show that in the uncertain demand environment the set of equilibria shrinks relative to the certain demand environment. They show that if the range of demand variation is unbounded then there is a unique SFE. If the range of demand variation is bounded then there is a continuum of equilibria, with the highest price corresponding to an equilibrium ranging from the competitive price to the Cournot price.

Green and Newbery (1992) applied the SFE model to analyze competition in the British wholesale electricity spot market. This market was run as a uniform price auction in which power sellers submit offer schedules. A supply function may be viewed as a continuous approximation of the discrete-unit offer schedules that are submitted in these auctions. Rather than assuming uncertain demand, Green and Newbery assume that demand varies deterministically over time during the course of a market day; deterministic variation in demand over time is mathematically equivalent to KM's model

possible bids, and symmetric costs and capacities.

of uncertain demand with bounded variation. Green and Newbery argue that in a symmetric model, suppliers should select the symmetric equilibrium that yields the highest profit. Using demand and cost parameters to reflect conditions in the England and Wales wholesale electricity market, they show that at the most profitable symmetric SFE, National Power and PowerGen, the two dominant bulk electricity suppliers, choose supply functions that yield prices far above marginal costs and cause substantial deadweight losses. Wolfram (1999) using actual pool outcomes shows that Green and Newbery's model does not describe the market very well, and the pool prices that they predict are much higher than the observed prices.

Newbery (1998) studies competition, contracts and entry in the electricity spot markets using analytically tractable models. He employs supply function type of strategy to model the spot market and a Cournot type strategy to model the contract market. He finds that, if the number of players increases, then maximum price reached in the pool and the average pool price fall. He also shows that, given that generators have capacity constraints, how a threat of entry would result in generators covering themselves with forward contracts, and this would yield more competition in the spot market, hence reducing average pool prices.

Baldick and Hogan (2002) study capacity constrained supply function equilibria and consider how they may be applied to electricity spot markets. Much of the SFE literature has focused on formulations in which firms have identical costs and capacities and in which firms play symmetric supply function strategies. Baldick and Hogan argue that asymmetries are common in electricty markets and that SFE models should take this into account. However, they show that the attempting to solve the system of differential equations that asymmetric SFE must satisfy is problematic. Solutions of the differential equations often fail the non-decreasing property, and so cannot be part of an equilibrium. Baldick and Hogan do not consider how the extent of excess capacity affects equilibrium predictions, nor do they consider the role that pivotal suppliers might play.

3. A Supply Function Equilibrium Model

In this section we formulate and explain the structures of the model we employ throughout the paper.

3.1 *Model Formulation*

We formulate a simple supply function equilibrium (SFE) model that highlights the role that pivotal suppliers play in determining the range of equilibrium outcomes. We assume that the quantity demanded is perfectly inelastic for prices up to some exogenous choke price, $\overline{p}$. This choke price could represent either buyers' maximum willingness to pay for wholesale electricity or, a price cap imposed by regulators. The load, or quantity demanded, in the market for prices up to $\overline{p}$ is given by the function $N(t)$. The load varies during the course of a market trading period (e.g., a day) according to the function $N(t)$, where $t \in [0,1]$ is an index of time during the trading period. If prices do not exceed $\overline{p}$, then total sales during the trading period are, $\int_0^1 N(t)dt$.

We assume a particular form for the load function in order to simplify calculations that we perform later. Our main qualitative results do not appear to be sensitive to the functional form.

(1) $N(t) = N(0)l^t$, where $l \in (0,1)$

The load during a trading period is ordered from its highest level to its lowest level according to the index t. The parameter l is referred to as the *load factor*. The load factor represents the ratio of the minimum load to the maximum load, during the market trading period.

We assume a very simple cost structure. There are $n \geq 2$ electricity suppliers. Each supplier has constant marginal cost of production, c, up to its capacity constraint. The capacity constraint for supplier i is K_i. Total capacity for all suppliers is, $K \equiv \sum_{i=1}^{n} K_i$. We assume throughout our analysis that $K \geq N(0)$; i.e., we assume that suppliers have sufficient capacity to meet the maximum demand (load). This "single step" cost function is quite restrictive. We employ it in order to simplify equilibrium calculations. Later in Section 7 we examine the implications of "multi-step" cost functions. We assume that $0 \leq c < \overline{p}$.

Each supplier i chooses a supply function, $S_i(p)$, which specifies the amount the supplier is willing to supply at each possible market price that might occur during a trading period. We require each supply function $S_i(p)$ to be continuous and non-decreasing in p, and to satisfy $S_i(p) \leq K_i$. Each supplier's supply function is held fixed during the trading period. A uniform price, $p(t)$, is established at each time $t \in [0,1]$ during the trading period. This price satisfies

(2) $$N(t) = \sum_{i=1}^{n} S_i(p(t)),$$

unless,

(3a) $$N(t) > \sum_{i=1}^{n} S_i(\overline{p})$$

in which case $p(t) = \overline{p}$ or,

(3b) $$N(t) < \sum_{i=1}^{n} S_i(0)$$

in which case $p(t) = 0$.[7]

3.2 *Supply Function Equilibrium*

A supply function equilibrium (SFE) is a Nash equilibrium in supply function strategies. Consider the profit for firm *i* at time *t*, given that its rivals have chosen supply functions, $\{S_j(p)\}_{j \neq i}$. If firm *i* chooses to supply the quantity $N(t) - \sum_{j \neq i} S_j(p(t))$ at time *t* then the market clearing price will be $p(t)$ and profit for firm *i* will be:

(4) $$\Pi_i(t) = (p(t) - c)[N(t) - \sum_{j \neq i} S_j(p(t))]$$

Firm *i* may also be viewed as choosing a price $p(t)$ a time *t* to maximize the profit expression in (4), where the expression in square brackets is the residual demand for firm *i*. If its rivals' supply functions are differentiable then the necessary condition for optimal price choice at time *t* for each firm *i* yields a system of ordinary differential equations for the *n* supply functions:

---

[7] If there are multiple prices that satisfy (2), then we take $p(t)$ to be the minimum price that satisfies (2).

$$(5) \qquad \sum_{j \neq i} S_j'(p) = \frac{S_i(p)}{(p-c)}, \; i = 1, \ldots, n$$

This is a simple version of the ODE system that has been intensively studied in the SFE literature.[8] We use equation (5) to derive candidates for SFE later in the paper.

SFE derived from (5) have the characteristic that a firm's supply function maximizes its profit at each point in time (t) during the trading period, given the supply functions chosen by its rivals. This implies that the way in which the load is distributed between its minimum and maximum levels (the load duration curve) does not change the set of SFE. In addition, if the maximum load is held fixed, changing the minimum load does not change the set of SFE; though it does change the predicted interval of prices observed in equilibrium.

3.3. *Pivotal Suppliers*

A supplier i is pivotal at time t during a market trading period if $\sum_{j \neq i} K_j < N(t)$. When this inequality holds, supplier i would be able to raise the market price at t all the way to $\overline{p}$ by either withholding some of its capacity or by submitting a supply function with some portion of its capacity bid in at price $\overline{p}$. Note that a pivotal supplier can move the market price to $\overline{p}$ unilaterally – that is, irrespective of the supply functions submitted by its rivals. If $\sum_{j \neq i} K_j < N(1)$ then supplier i is pivotal for the entire trading period.

[8] This version is simple because there is no demand function term, due to perfectly inelastic demand, and because marginal cost is constant.

We emphasize that there are three important parameters in our analysis, which we treat as exogenous. One parameter is the number of suppliers, $n \geq 2$. A second is the load factor, l. A third parameter is the capacity index, k. The capacity index is the ratio of the total amount of production capacity, K, held by all n suppliers, to the maximum load, $N(0)$; i.e., $k \equiv K / N(0)$. Our assumption regarding K implies that $k \geq 1$. Larger values of k indicate greater excess capacity held by suppliers.

4. Symmetric Firms Model

In the previous section it was assumed that all suppliers have a common marginal cost c for production up to capacity. In this section we also suppose that total capacity is equally divided among suppliers; $K_i = K / n,\ i = 1,...,n$. For this symmetric suppliers formulation, equation (5) may be used to derive a continuum of candidate symmetric SFE. We use the term candidate because some symmetric supply functions derived from (5) may fail to be equilibria due to the presence of pivotal suppliers.

If all suppliers utilize a common supply function, $S(p)$, then the system of equations in (5) simplify to the following single ODE:

$$(6) \qquad S'(p) = \frac{S(p)}{(n-1)(p-c)}$$

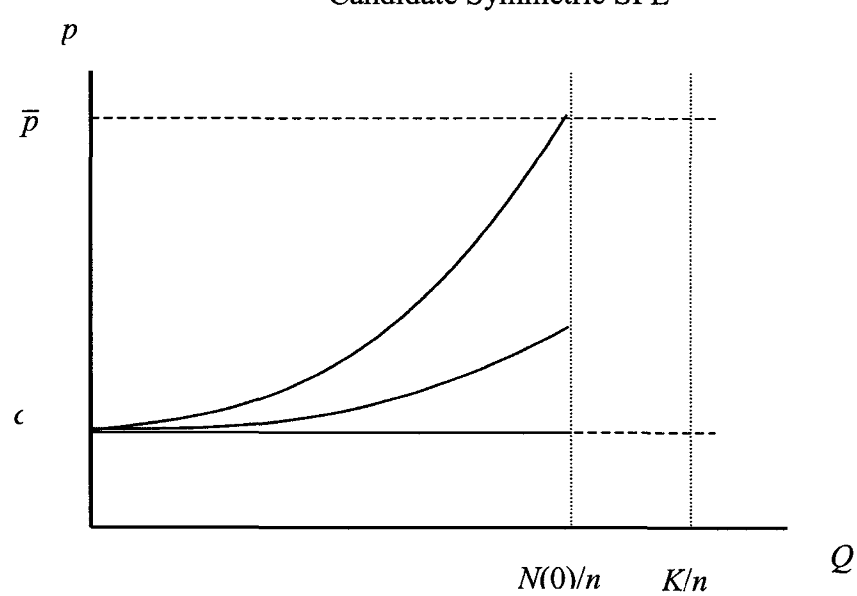
There is a continuum of solutions to (6) of the form:

$$(7) \qquad S(p) = \frac{N(0)}{n}\left[\frac{p-c}{p(0)-c}\right]^{1/(n-1)}$$

The supply function solutions in (7) are a special case of results in Rudkevich, et al (1998). They derive SFE solutions for a symmetric model with zero demand elasticity and general cost functions (including multi-step marginal cost functions).

The supply functions in (7) are indexed by the initial price, $p(0)$, which can take on any value in the interval, $(c, \overline{p}]$. $p(0)$ is the time zero price for the load ordering we are using. If $n = 2$ then the supply functions in (7) are linear. If $n > 2$ then the supply functions are convex, as in Figure One. This figure illustrates several of these solutions, including the limiting case of the most competitive solution (a horizontal supply function at $p = c$) and the least competitive solution (with $p(0) = \overline{p}$).

Figure One
Candidate Symmetric SFE



Each candidate SFE in (7) is indexed by an initial price $p(0) \in (c, \bar{p}]$. Given the functional forms we have assumed for demand, cost, and the load distribution, it is straightforward to calculate the equilibrium prices and profit per firm, for each initial price and associated candidate SFE.

$$p(t) = c + (p(0) - c)l^{t(n-1)}$$

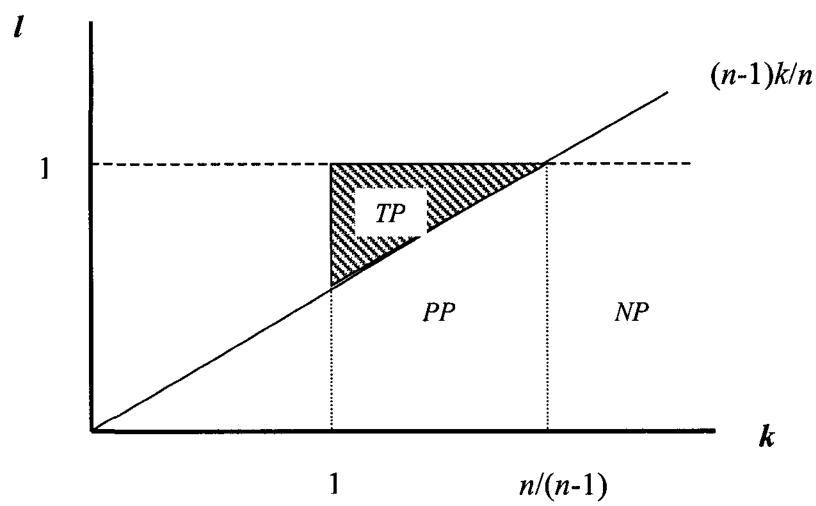
$$\text{(8)} \qquad \Pi^{SFE} = \int_{t=0}^{1} (p(t) - c)(N(t)/n)dt = \frac{N(0)(p(0) - c)(1 - l^{n})}{-n^{2} \ln l}$$

Π^{SFE} is profit per firm associated with a candidate SFE, where it is understood that this profit depends on the initial price $p(0)$ that indexes the candidate SFE. This profit depends on parameters such as the load factor, the number of firms, marginal cost, and maximum load, as well as on the initial price. We note that this profit is independent of the capacity index, k.

Next we examine how the presence of pivotal suppliers and the extent of excess capacity influence the set of equilibria. Figure Two illustrates relevant parameter ranges for the capacity index and the load factor, for a fixed number of suppliers, n. The load factor, l, lies between zero and one. The capacity index, k, is greater than or equal to one. If $l \in [k(n-1)/n, 1)$ then each firm is pivotal at all times t during the trading period. We refer to this as the totally pivotal (TP) case. Formally, we define the set, $TP \equiv \{(k, l, n) : n \geq 2, n/(n-1) \geq k \geq 1, l \in [(n-1)k/n, 1)\}$. This set is illustrated by the shaded area in Figure Two. A second situation arises when each firm is pivotal for some times, but not all times during the trading period. We refer to this as the partially pivotal (PP) case. The parameter set PP is directly below set TP in Figure Two;

$PP \equiv \{(k,l,n): n \geq 2, n/(n-1) \geq k \geq 1, l \in (0,(n-1)k/n)\}$. If parameters are in set *PP* then each firm is pivotal for times between zero and τ satisfying, $N(\tau) = (n-1)K/n$. A third case occurs when no firm is pivotal at any time; we refer to this as the never pivotal (NP) case. This occurs when $k \geq n/(n-1)$, and is illustrated by the area *NP* in Figure Two. In this case, any collection of $n-1$ suppliers has enough capacity to serve the maximum load.

Figure Two
Parameter Ranges for Capacity Index (k)
and Load Factor (l)



If $k \geq n/(n-1)$, then parameters are in the *NP* (never pivotal) region. In this case, supply functions in (7) are SFE for all initial prices, $p(0) \in (c, \bar{p}]$. The set of equilibria includes supply functions that are arbitrarily close to the competitive equilibrium of pricing at marginal cost at all times during the trading period. These supply functions can be shown to be equilibrium strategies by using the approach proposed by Klemperer and Meyer (1989, p. 1255). For any candidate SFE with initial price $p(0) \in (c, \overline{p}]$, extend the supply function linearly for prices above $p(0)$ with slope, $S'(p(0))$, until the quantity supplied reaches capacity, K/n. Then at each time *t*, $p(t)$ and $S(p(t))$ are the globally optimal price and quantity, given rivals' (extended) supply functions.

If parameters are in the *TP* or *PP* regions then an individual firm has an incentive to deviate from $S(p)$ defined in (7) for some initial prices in the interval $(c, \overline{p}]$. We offer an informal argument for this initially, in order to develop the intuition. Following this, we provide a formal argument with propositions that delineate the set of equilibria.

For our informal argument, suppose that $p(0) = c + \varepsilon$, for some small, positive ε. The profit Π^{SFE} that a firm earns by following this supply function strategy is proportional to ε. It remains true that at each time *t*, $p(t)$ and $S(p(t))$ are locally optimal price and quantity choices, given rivals' supply functions. However, each firm is pivotal for at least some times during the trading period. By bidding in some of their supply at a price equal to $\overline{p}$ at a time *t* when they are pivotal, a deviating firm has sales equal to $N(t) - (n-1)K/n$ and earns profit, $(\overline{p} - c)[N(t) - (n-1)K/n]$. Integrating these profits over all times for which the firm is pivotal will yield total profit that exceeds Π^{SFE} if ε is small. $p(t)$ and $S(p(t))$ are locally optimal price and quantity choices at time *t*, but they are not globally optimal choices at *t* due to the capacity constraints of rival firms.

Next we move to a more formal argument regarding the role that pivotal firms play in restricting the equilibrium set. We will examine a simple type of deviation in this section. The deviation involves a supply function with no units offered for prices below the maximum price, $\bar{p}$, and all units up to capacity offered at price $\bar{p}$.[9] We will compare the profit associated with this simple type of deviation with profit from candidate SFE's. In Section 5 we consider optimal deviations.

Suppose that parameters are in the totally pivotal region; $(k,l,n) \in TP$. Then the residual demand at time t for the deviating firm at price $\bar{p}$ is, $[N(t)-(n-1)K/n] > 0$. Total profit for the deviating firm is,

(9) $$\Pi^D = \int_0^1 [N(t)-(n-1)K/n](\bar{p}-c)\,dt = \frac{N(0)(\bar{p}-c)[n^2(1-l)+n(n-1)k\ln l]}{-n^2 \ln l}.$$

Deviation profit exceeds profit associated with the candidate SFE if $\Pi^{SFE} < \Pi^D$, or equivalently if,

(10) $$p(0)-c < \phi(k,l,n)(\bar{p}-c)$$

where,

(11) $$\phi(k,l,n) \equiv [n^2(1-l)+n(n-1)k\ln l]/[1-l^n].$$

The function ϕ produces a fraction for each triple (k,l,n) of parameters in the set *TP*. If the markup at the initial price for a candidate SFE is less than the fraction ϕ of the maximum markup, $\bar{p}-c$, then the candidate SFE is not an equilibrium. For example, if

[9] If the firm is required to submit a supply function that is single-valued at each price, then the deviation may be modified so that no units are offered for prices below $\bar{p}$-delta for some small, positive delta, with a sharply increasing supply for prices in the interval [$\bar{p}$-delta, $\bar{p}$].

the capacity index is k =1.2, the load factor is $l = 0.85$, and the number of firms is $n = 3$, then $\phi(k,l,n) \approx 0.46$. Any candidate SFE that does not involve a markup at the initial price that is at least 46 % of the maximum markup is ruled out as an equilibrium.

Proposition 1: The ϕ function is decreasing in the capacity index k, increasing in the load factor l, and decreasing in the number of suppliers n for $(k,l,n) \in TP$.

Proof – see appendix.

Equation (10) provides a sufficient condition for eliminating certain candidate SFE as equilibria. Proposition 1 shows that the set of candidate supply functions that are ruled out becomes larger as the amount of excess capacity decreases, as the load factor increases, and as the number of firms (holding total capacity constant) decreases.

Corollary 1: The value of $\phi \to 1$ as $k \downarrow 1$ and $l \uparrow 1$.

Proof: This result may be demonstrated by setting $k = 1$ and applying l'Hopital's Rule to evaluate $\lim_{l \uparrow 1} \phi(1,l,n)$.□

As the amount of excess capacity goes to zero and the load factor approaches 100%, the set of equilibria shrinks to the least competitive SFE; the SFE with initial price $p(0) = \overline{p}$. Note that the limiting result in Corollary 1 holds regardless of the number of firms that capacity is divided among.

Several papers in the SFE literature have emphasized how binding capacity constraints limit the set of equilibria. For example, see Green and Newbery (1992) and

Baldick and Hogan (2002). The approach that these papers appear to take is to rule out a solution to the differential equations analogous to (5) as a SFE whenever the solution violates a capacity constraint for at least one firm along the equilibrium path of prices. It is important to note that the solutions we are eliminating as equilibria via equation (10) are not ruled out because they violate capacity constraints. All supply functions defined by (7) with $p(0) \in (c, \overline{p}]$ satisfy capacity constraints along the candidate equilibrium path of prices. Equilibria are ruled out because capacity constraints limit the ability of rival firms to expand output when any one firm pushes up prices by limiting its own output (as offered via its supply function), even though capacity constraints are *non-binding* along symmetric candidate supply function equilibria.

Now suppose that parameters are in the partially pivotal region; $(k,l,n) \in PP$. We continue to suppose that a deviating firm offers no units for prices below $\overline{p}$ and as many units as demanded at the price ceiling, $\overline{p}$. Residual demand at time t for the deviating firm at price $\overline{p}$ is, $[N(t)-(n-1)K/n] > 0$, if $t < \tau \equiv \ln[(n-1)k/n]/\ln(l)$ and residual demand is zero for $t \geq \tau$. Total profit for the deviating firm is,

$$(12) \qquad \hat{\Pi}^D = \int_0^{\tau} [N(t)-(n-1)K/n](\overline{p}-c)\,dt = \frac{N(0)(\overline{p}-c)[n^2(1-l^{\tau})+n(n-1)k\tau \ln l]}{-n^2 \ln l},$$

where τ defined above depends on (k,l,n).

Deviation profit exceeds profit associated with the candidate SFE if $\Pi^{SFE} < \hat{\Pi}^D$, or equivalently if,

$$(13) \qquad p(0)-c < \lambda(k,l,n)(\overline{p}-c)$$

where,

(14) $\lambda(k,l,n) \equiv [n^2(1-l^\tau)+n(n-1)k\tau \ln l]/[1-l^n]$, with $\tau \equiv \ln[(n-1)k/n]/\ln(l)$.

Equations (13) and (14) provide a sufficient condition that eliminates certain candidate SFE as equilibria for the partially pivotal case. If the markup at the initial price for a candidate SFE is less than the fraction λ of the maximum markup, $\bar{p}-c$, then the candidate SFE is not an equilibrium. The following proposition is analogous to Proposition 1.

Proposition 2: The λ function is decreasing in the capacity index *k*, increasing in the load factor *l*, and decreasing in the number of suppliers *n* for $(k,l,n) \in PP$.

Proof – see appendix.

5. Optimal Deviations in the Symmetric Firms Model

In the previous section we showed how some candidate SFE could be ruled out as equilibria via a simple kind of deviating supply schedule when there are pivotal firms. The deviation we examined involved bidding all units in at the maximum acceptable price, $\bar{p}$. However, this kind of deviation may not be an optimal response to rivals' supply functions. In this section we characterize optimal deviations from candidate SFE. Consideration of optimal deviations expands the set of candidate symmetric SFE that are ruled out as equilibria.

We form an optimal control problem to be able to characterize optimal deviations as follows:

Let $t \in [0,1]$, $p(t)$ be the state variable, and $u(t)$ be the control variable. Also let residual demand for a deviating firm be $D^R(p(t),t)$, which is defined as

$$D^R(p(t),t) = \begin{cases} N(t) - \sum_{j \neq i} K_j , & if \quad p > p_0 \\ N(t) - \sum_{j \neq i} S_j(p) , & if \quad p \leq p_0 . \end{cases}$$

Then the control problem becomes;

$$\max_{\{u(t)\}_{t=0}^{1}} J(.) = \int_0^1 \left[p(t) D^R(p(t),t) - C(D^R(p(t),t)) \right] dt$$

$$\begin{aligned} s.t. \quad & (15) \quad p'(t) = u(t) \\ & (16) \quad p(0) = \bar{p} \\ & (17) \quad D_1^R(p(t),t)u(t) + D_2^R(p(t),t) \leq 0 \\ & (18) \quad D_1^R(p(t),t)u(t) + D_2^R(p(t),t) - \delta u(t) \geq 0 \end{aligned}$$

where, $D_1^R(p(t),t) \equiv \dfrac{\partial D^R(.)}{\partial p}$, $D_2^R(p(t),t) \equiv \dfrac{\partial D^R(.)}{\partial t}$, and δ is a positive number. The Objective function denotes a deviating firm's profit. Inequality (17) is due to definition of supply function and market price. That is, let $S(p(t))$ be the supply function chosen by a firm. Then the market-clearing price satisfies $D^R(p(t),t) - S(p(t)) = 0$. If we differentiate this with respect to *t*, then $D_1^R(.)p'(.) - D_2^R(.) = S'(.)p'(.)$ holds. Since the supply function is required to be non-decreasing and $p'(.) < 0$, we obtain constraint (17). The last constraint (18) puts an upper bound on the rate of change of quantity supplied with respect to price.

Now let us define the *Hamiltonian* of the above problem as

$H(.) = p(t)D^R(p(t),t) - C(D^R(p(t),t)) + \lambda(t)u(t) - \theta(t)F(p(t),u(t),t) + \beta(t)G(p(t),u(t),t)$,

where $\lambda(t)$, $\theta(t)$, $\beta(t)$ are the multiplier functions of the first, third and forth constraints respectively. $F(.)$ and $G(.)$ denote left hand side values of the constraints (17) and (18) respectively.

Necessary conditions of the above maximization problem are the constraints (15) - (18) and

(19) $\quad \partial H/\partial u = \lambda(t) - \theta(t)D_1^R(.) + \beta(t)[D_1^R(.) - \delta] = 0$

(20) $\quad \lambda'(t) = -\partial H/\partial p = -\left[D^R(.) + D_1^R(.)[p(t) - C'(.)] + (\beta(t) - \theta(t))(D_{11}^R(.)u(t) + D_{21}^R(.))\right]$

(21) $\quad \theta(t) \geq 0\,,\ \ \theta(t)F(.) = 0$

(22) $\quad \beta(t) \geq 0\,,\ \ \beta(t)G(.) = 0$

Note that $D_{21}^R(.) \equiv \partial D_2^R/\partial p = 0$.

Now let us piece together the features of an optimal solution.

Free intervals. At time t in a time interval for which (17) and (18) are non-binding, we have $\theta(t) = 0$, and $\beta(t) = 0$. Then by (19), $\lambda(t) = 0$, which implies $\lambda'(t) = 0$. Hence by (20), we obtain $0 = D^R(.) + D_1^R(.)[p(t) - C'(.)]$, which is the basic SFE condition if price is in the 'SFE range', that is if $p(t) < p^0$ holds, where p^0 is the SFE initial price. However for some time t if $p(t) > p^0$ then the firm is "defecting" to a higher price, that is not a part of the SFE. The following is about this case.

Constrained Intervals. First we consider the case in which (17) binds. In this case for some time t we must have $D^R(p(t),t)$ equal to a constant, and

$p'(t) = u(t) = \dfrac{-D_2^R(p(t),t)}{D_1^R(p(t),t)} < 0$. That is $p(t)$ must be falling at precisely the rate that leaves the residual demand quantity constant. Second we consider the case in which (18) binds. Then for such a time t $D_1^R(.)u(t) + D_1^R(.) = \delta u(t)$, and $p'(t) = \dfrac{-D_2^R(.)}{D_1^R(.) - \delta} < 0$ holds. This can happen only during an initial time interval during which the derivative of the firm's supply function with respect to price is constrained to be less than or equal to δ.

Now let us keep the same assumptions we considered in the previous section regarding load, cost, capacity etc. Given these assumptions for free and constrained intervals we can write a deviating firm's maximized profit as

$$\begin{aligned} \Pi^{opt}(.) = & \int_0^{t_1} \left(N(t) - (n-1)K/n\right)\left[\left(N(t) - N(0) + \delta\bar{p}\right)\delta^{-1} - c\right]dt \\ & + \int_{t_1}^{t_2} \left(p(0) - c\right)\tilde{q}\,dt + \int_{t_2}^{t_3} (p(t) - c)\tilde{q}dt + \int_{t_3}^{1} \left(p(t) - c\right)N(t)n^{-1}\,dt \end{aligned} \quad (23)$$

The first integral considers the constrained interval in which (18) binds. The first term in the first integrand is the residual demand for a deviating firm. The second term is the price at this time t minus marginal cost. This price can be computed given above necessary conditions along with the price at time zero equal to "choke price", $\bar{p}$, and the quantity supplied at this time equals to $N(0) - (n-1)K/n$. The second integral considers the case in which (17) binds so that quantity supplied ($\tilde{q}$) will be constant in the time interval (t_1, t_2). Here $\tilde{q} = N(t_1) - (n-1)K/n$ and $\tilde{q} = N(t_2) - (n-1)N(0)/n$ hold. The third integral is the special case of the constraint (17) in which quantity supplied, $\tilde{q}$, is constant and price during this period satisfies $N(t)-\tilde{q}=\sum_{j\neq i}S_j(p(t))$, and t_3 satisfies $N(t_3)=n\tilde{q}$. The last integral is the case for free intervals, where the integral reflects the profit for SFE in the time interval $(t_3,1]$, and here of course $p(t)$ is the SFE price. Consequently the price keeps decreasing till time t_1, at this time the price jumps down to p^0. Between t_1 and t_2 the price stays at p^0 level and after t_2 till time 1 it keeps decreasing.

One can easily take the above integrations and report the profit function. For the sake of brevity we omit that, since resultant function is long. However we note that this profit is a function of the vector of parameters $\left(N(0),\overline{p},c,n,k,l,\delta,p^0,\tilde{q}\right)$. One can also perform a "grid search" to find the optimal values of t_1, t_2, t_3 for this optimal deviation problem. Basically the optimal t_1, which is the time after which quantity supplied becomes constant, satisfies $t_1\in\arg\max\left\{\Pi^{opt}(.)\right\}$. After obtaining optimal t_1, from the relations $N(t_2)=\tilde{q}+(n-1)N(0)/n$, and $N(t_3)=n\tilde{q}$ one can also obtain optimal t_2, t_3.

The following example considers a case in which optimal deviation profit exceeds the simple kind of deviation profit that we examined in the previous section.

<u>*Example 1*</u>: Note that simple kind of deviation profit in the partially pivotal region can be calculated as

$\hat{\Pi}^D = \int_0^{\tau}[N(t)-(n-1)K/n](\bar{p}-c)\,dt = \dfrac{N(0)(\bar{p}-c)[n^2(1-l^{\tau})+n(n-1)k\tau\ln l]}{-n^2\ln l}$. We above

denoted optimal deviation profit as $\Pi^{opt}(.)$. Now let number of players be $n=2$, capacity index be $k=1.01$, load factor be $l=0.04$, total capacity be $K=940$, choke price or price cap be $\bar{p}=17.5$, peak load price be $p^0=15.4$, and marginal cost be $c=12$. Given these values, we perform grid search to find optimal switching times, the constant quantity supplied from t_1 to t_3, and the corresponding profits. We obtain that as $\delta\to\infty$, $t_1=0.073$, $\tilde{q}=266.3$, $t_2=0.075$, $t_3=0.173$. We also obtain SFE profit, $\Pi^{SFE}=245.37$, and simple kind of deviation profit in the *PP*, $\hat{\Pi}^D=238.5$, for which $\tau=0.212$. However, optimal deviation profit becomes $\Pi^{opt}=320.46$.

In this example we showed that a simple deviation does not rule out a candidate SFE with initial price $p^0=15.4$. However, by means of the optimal deviation, SFE at that price is being ruled out.

6. Asymmetric Firms Model

In reality firms are asymmetric in terms of capacities and cost functions especially in the electric industry. Concerns about market power of pivotal suppliers in wholesale electricity markets have often focused on the largest suppliers (e.g. see Lave and Perekhodtsev (2001)). Suppliers with the greatest generation capacity are the ones most likely to be able to force up the market price by withholding production. However there is relatively little analysis of models in which firms have asymmetric capacities in the supply function literature. Green and Newbery (1992) discuss equilibria for asymmetric duopoly. However they say that it is, "... more difficult to solve for the pair of (differential) equations for the asymmetric equilibrium than the single equation for the symmetric equilibrium, and the rest of the paper will restrict attention to the symmetric case." Baldick and Hogan (2002) also mention asymmetric firms in capacities and cost functions. However they say that the differential equation approach of solving supply functions may not be effective, because the resulting supply functions may fail the non-decreasing property.

In this section we provide an analytic solution of supply function equilibria of asymmetric firms in capacities. Because of the relative simplicity of the model, we can provide a precise condition under which supply function solutions to the differential equations satisfy the non-decreasing property. We show that the larger firm is more pivotal than the smaller firm in both totally and partially pivotal regions. The larger firm has an incentive to deviate from a wider range of SFE, and it is the larger firm's deviation incentives that determine which candidate SFE are ruled out as equilibrium.

Here, in particular, we focus on duopoly, in which each firm is only different in capacities. We assume that, without loss of generality, firm 2 has less capacity than its competitor. We also assume throughout our analysis that $K_1 + K_2 \geq N(0)$. If $K_1 > K_2 \geq \frac{1}{2}N(0)$ then the symmetric SFE that were characterized in Section 4 are feasible, since each firm has enough capacity to meet at least one-half of the peak demand. In this case the incentives to defect from any candidate symmetric SFE are driven by the amount of capacity of the smaller firm. The larger firm's incentive to deviate are identical to the incentives of a symmetric firm in a duopoly market with capacity index $k = 2K_2 / N(0)$. Using the load factor, l, $k = 2K_2 / N(0)$, and $n = 2$, the ϕ and λ functions derived in Section 4 define the sets of equilibria ruled out by simple defections.

If $K_1 > \frac{1}{2}N(0) > K_2$ then the symmetric SFE are not feasible because they violate the capacity constraint for firm 2.

If we follow the analysis we did in subsection 3.2, we obtain following differential equations for the supply functions:

$$(24) \qquad S_2' = \frac{S_1(p)}{(p-c)}, \quad S_1' = \frac{S_2(p)}{(p-c)},$$

where S_1 and S_2 are the supply functions for firm 1 and firm 2 respectively. If we differentiate the second equation one more time, we obtain $S_1'' = \frac{(p-c)S_2' - S_2}{(p-c)}$. Now plug above two equations into this equation, and then get the following homogeneous second order linear differential equation,

$$(25) \qquad S_1'' = \frac{S_1}{(p-c)^2} - \frac{S_1'}{(p-c)}.$$

Boundary conditions are $S_2(p(0)) = K_2$. It is possible that the smaller firm could have even smaller initial output, coupled with more output for the larger firm, but this decreases firm 2's profit. Here we assume that smaller firm plays aggressively so that he supplies his all available capacity at time zero. Then $S_1'(p(0)) = \dfrac{K_2}{p(0)-c}$ holds. Now, for firm 1 we have a homogeneous second order linear differential equation with two boundary conditions ($S_1(p(0))$, $S_1'(p(0))$). Solution of this is the supply function for firm 1:

$$(26) \qquad S_1(p) = \frac{N(0)}{2}\left[\frac{p-c}{p(0)-c} + \frac{p(0)-c}{p-c}\right] - K_2\left[\frac{p(0)-c}{p-c}\right].$$

By performing similar analysis, we obtain the following supply function for firm 2:

$$(27) \qquad S_2(p) = \frac{N(0)}{2}\left[\frac{p-c}{p(0)-c} - \frac{p(0)-c}{p-c}\right] + K_2\left[\frac{p(0)-c}{p-c}\right].$$

Note that $S_1(p) + S_2(p) = N(0)\dfrac{p-c}{p(0)-c}$. Also, if $S_1(p(0)) = S_2(p(0))$, then the symmetric supply function equilibria are obtained.

Now we can calculate the equilibrium prices from the following relation:

$$(28) \qquad N(t) = S_1(p(t)) + S_2(p(t)) = N(0)\frac{p(t)-c}{p(0)-c},$$

which implies that $p(t) = c + \dfrac{N(t)}{N(0)}(p(0)-c)$. Let $k_2 \equiv \dfrac{K_2}{N(0)}$, then profit for firm 1 becomes

$$(29) \qquad \Pi_1^{SFE} = \int_{t=0}^{1} (p(t)-c)S_1(p(t))dt = \frac{(p(0)-c)N(0)}{2}\left[\frac{l^2-1}{2\ln l} + 1 - 2k_2\right].$$

Note that this profit is dependent on small firm's capacity index k_2, as well as the initial price, maximum load, load factor, and marginal cost.

Similarly one can calculate firm 2's profit as

$$\text{(30)} \qquad \Pi_2^{SFE} = \int_{t=0}^{1} (p(t)-c)S_2(p(t))dt = \frac{(p(0)-c)N(0)}{2}\left[\frac{l^2-1}{2\ln l} - 1 + 2k_2\right].$$

Observe that summation of the profits of firm 1 and firm 2 under SFE gives the total profit earned under the symmetric SFE with the same initial price.

Each firm's supply function should be non-decreasing in price in order to be part of the equilibrium. If the first derivative of firm 1's supply function becomes zero then $p-c=\sqrt{1-2k_2}\left(p(0)-c\right)$ holds. Inserting the equilibrium prices provided above into the left hand side of this equation, we obtain $l=(1-2k_2)^{1/2t}$. By using the boundary conditions we rewrite the minimum load factor in terms of $S_1\left(p(0)\right)$ and $S_2\left(p(0)\right)$ as $\tilde{l} \equiv \left(\frac{S_1\left(p(0)\right)-S_2\left(p(0)\right)}{S_1\left(p(0)\right)+S_2\left(p(0)\right)}\right)^{1/2} = (1-2k_2)^{1/2}$. Thus, the load factor should exceed this threshold $\tilde{l}$ to ensure that asymmetric supply function equilibria exist. A load factor in excess of this threshold also insures that the small firm's supply quantity is non-negative along the equilibrium path.

Next as we did in section 4, we examine how the presence of pivotal suppliers and the extent of excess capacity influence the set of equilibria. In this asymmetric case, the totally pivotal (TP) set is defined as $TP_{asym} \equiv \{(k_2,l) \mid 0<k_2<0.5,\ \max\{k_2,\sqrt{1-2k_2}\ \} \le l < 1\}$. Under this circumstance, the residual demand at time t for the deviating firm 1 at price $\bar{p}$ is, $(N(t)-K_2)>0$. Then his profit is calculated as

(31) $$\Pi_1^D = \int_{t=0}^{1} (\bar{p}-c)(N(t)-K_2)dt = N(0)(\bar{p}-c)\left[\frac{l-1}{\ln l}-k_2\right].$$

It will also be necessary to consider the deviation incentives for the smaller firm. One can obtain firm 2's profit in that region as follows;

(32) $$\Pi_2^D = \int_{t=0}^{1} (\bar{p}-c)(N(t)-K_1)dt = N(0)(\bar{p}-c)\left[\frac{l-1}{\ln l}-k_1\right], \text{ where } k_1 \equiv K_1/N(0).$$

For firm 1, deviation profit exceeds profit associated with the candidate SFE if $\Pi_1^D > \Pi_1^{SFE}$, or equivalently if,

(33) $$p(0)-c < \psi_1(k_2,l)(\bar{p}-c)$$

where,

(34) $$\psi_1(k_2,l) \equiv \frac{4(l-1-k_2\ln l)}{(l^2-1+2\ln l-4k_2\ln l)}.$$

This function, ψ_1, generates a fraction for each double (k_2,l) of the parameters in the set TP_{asym}. If the markup at the initial price for a candidate SFE is less than the fraction ψ of the maximum markup, $\bar{p}-c$, then the candidate SFE is not an equilibrium.

One can similarly derive the relationship between firm 2's profit under SFE and his profit under TP_{asym} as $\Pi_2^D > \Pi_2^{SFE}$, or equivalently

(35) $$p(0)-c < \psi_2(k_1,k_2,l)(\bar{p}-c),$$

where ,

(36) $$\psi_2(k_1,k_2,l) \equiv \frac{4(l-1-k_1\ln l)}{(l^2-1-2\ln l+4k_2\ln l)}.$$

The following proposition sets the relationship between ψ_1 and ψ_2, and shows that ψ_2 eliminates fewer candidate SFE than ψ_1 does.

Proposition 3: The firm with the larger share of capacity has a greater incentive to deviate from a supply function equilibrium. The larger firm's payoffs from a simple deviation rule out a larger set of candidate SFE than do the smaller firm's payoffs from a simple deviation. That is, $\psi_2 < \psi_1$.

Proof – see appendix.

Proposition 3 indicates that the larger firm is more pivotal than the smaller firm. The larger firm has an incentive to deviate from a wider range of SFE, and it is the larger firm's deviation incentives that determine which candidate SFE are ruled out as equilibrium. The following proposition characterizes ψ_1.

Proposition 4: The ψ_1 function is increasing in the load factor *l*, and decreasing in the (firm 2's) capacity index k_2 for $(k_2, l) \in TP_{asym}$.

Proof – see appendix.

Similar to equation (10), equation (33) provides a sufficient condition for eliminating some set of SFE as equilibria. Proposition 4 means that the set of equilibria being ruled out becomes larger as capacity index decreases, and the load factor increases.

Corollary 2: The value of $\psi_1 \to 1$ as $k_2 \downarrow 0$ and $l \uparrow 1$.

Proof: The proof is similar to the proof of Corollary 1. □

Next we will show for an example how the set of equilibria ruled out vary with the small firm's share of capacity. Before we introduce this example we can easily show that

$\phi(l,n=2,k=1)=\psi_1(l,k_2=0.5)=\psi_2(l,k_2=0.5,k_1=0.5)$, where variables are defined in their domains. We just insert the values in to the functions, and obtain

$$\phi(l,k,n)=\left.\frac{n^2(1-l)+n(n-1)k\ln l}{1-l^n}\right|_{n=2,k=1}=\frac{4(1-l)+2\ln l}{1-l^2},$$

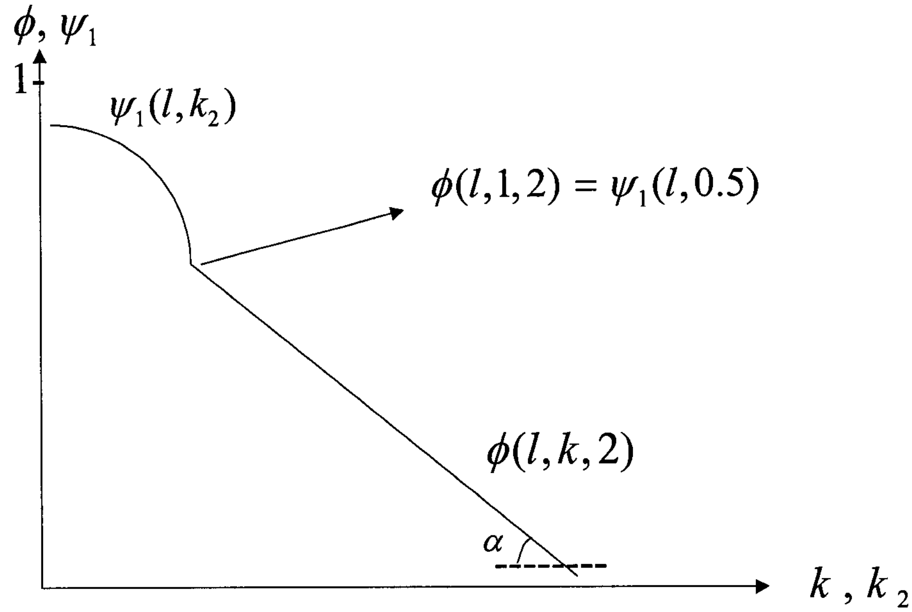
$$\psi_1(l,k_2)=\left.\frac{4(l-1-k_2\ln l)}{l^2-1+2\ln l-4k_2\ln l}\right|_{k_2=0.5}=\frac{4(l-1-0.5\ln l)}{l^2-1},$$

$$\psi_2(l,k_2,k_1)=\left.\frac{4(l-1-k_1\ln l)}{l^2-1-2\ln l+4k_2\ln l}\right|_{k_2=0.5=k_1}=\frac{4(l-1-0.5\ln l)}{l^2-1}.$$

The following example is a case of this relation, and we see how these functions behave with respect to the share of total capacity.

<u>*Example 2*</u>: Let us say load factor is $l=0.9$. In this case the capacity index of small firm in the totally pivotal region must be in the interval $0.095\le k_2<0.5$. Also symmetric case capacity index in the totally pivotal set should satisfy $1\le k\le 1.8$. In these intervals, we plot the factor functions in Figure Three.

Figure Three
Factor functions with the share of total capacity



Here the slope of the function ϕ is -1.109. It shows the rate of change of ruled out candidate SFE equilibria with respect to change in total capacity index. The ϕ function follows linear path and intersects at $\phi(l,1,2) = \psi_1(l,0.5)$, after this point ψ_1 follows concave path, and its value at $k_2 = 0.095$ takes a number close to 1.

In this example Figure Three characterizes how the rate of candidate SFE are being ruled out with respect to small firm's capacity share for a fixed load factor in which players are totally pivotal. If players have the same capacity then candidate SFE is being ruled out linearly with respect to capacity change. However, if the players become asymmetric and the small firm's share decreases then candidate SFE are being ruled out logarithmically.

In the asymmetric case another situation arises when each firm is pivotal for some trading period. In this case the asymmetric partially pivotal (PP_{asym}) region is defined as the set $PP_{asym} \equiv \{(k_2,l) \mid 0 < k_2 < 0.5, \sqrt{1-2k_2} \le l < k_2\}$, where each firm is pivotal for times between zero and $\hat{\tau}$ satisfying, $N(\hat{\tau}) = K_2$. In particular, $\hat{\tau} = \ln(k_2)/\ln l$. Note that $PP_{asym} \equiv \varnothing$ holds, if $k_2 < \sqrt{1-2k_2}$, or $k_2 < 0.414$. Given that parameters are in this set, residual demand at time t for deviating firm 1 at price $\bar{p}$ is, $(N(t)-K_2) > 0$, if $t < \hat{\tau}$, and residual demand is zero for $t \ge \hat{\tau}$. Then profit for the deviating firm 1 is,

(37) $$\hat{\Pi}_1^D = \int_{t=0}^{\hat{\tau}} (\bar{p}-c)(N(t)-K_2)dt = N(0)(\bar{p}-c)\left[\frac{l^{\hat{\tau}}-1}{\ln l} - \hat{\tau}k_2\right].$$

Similarly we obtained deviating firm 2's profit in this set as

(38) $$\hat{\Pi}_2^D = \int_{t=0}^{\hat{\tau}} (\bar{p}-c)(N(t)-K_1)dt = N(0)(\bar{p}-c)\left[\frac{l^{\hat{\tau}}-1}{\ln l} - \hat{\tau}k_1\right].$$

For firm 1, deviation profit exceeds profit associated with the candidate SFE if $\hat{\Pi}_1^D > \Pi_1^{SFE}$, or equivalently if,

(39) $$p(0) - c < \beta_1(k_2,l)(\bar{p}-c)$$

where,

(40) $$\beta_1(k_2,l) = \frac{4(l^{\hat{\tau}} - 1 - \hat{\tau}k_2 \ln l)}{(l^2 - 1 + 2\ln l - 4k_2 \ln l)}.$$

The β_1 function generates a fraction for each double (k_2,l) of the parameters in the set PP_{asym}. Again, if the markup at the initial price for a candidate SFE is less than the fraction β_1 of the maximum markup, $\bar{p}-c$, then the candidate SFE is not an equilibrium.

The following proposition and corollaries characterize the function β_1.

Proposition 5: The β_1 function is decreasing in the (firm 2's) capacity index k_2, and increasing in the load factor l for $(k_2, l) \in PP_{asym}$.

Proof – see appendix.

Similar to equation (33), equation (39) provides a sufficient condition for eliminating some set of SFE as equilibria. Proposition 5 means that the set of equilibria being ruled out becomes larger as capacity index of player 2 decreases, and the load factor increases.

Corollary 4: The value of $\beta_1 \to 0$ as $k_2 \downarrow 0$ and $l \uparrow k_2$.

Proof: The proof is very similar to the proof of Corollary 1. □

Corollary 5: $\hat{\tau} > \beta_1$.

Proof – see appendix.

Corollary 5 indicates that as trading period tightens, that is it approaches to zero, the peak load also decreases, and the deviating firm's profit decreases, hence the larger firm will have less incentive to deviate, thus fewer candidate SFE will be ruled out as equilbria. It also shows that firm 1's deviation rate is always less than the maximum allowed trading period in which deviation is possible.

We have also studied firm 2's fraction function that sets relationship between SFE profit and deviation profit in the partially pivotal region. We have found that firm 1 is more pivotal than firm 2 for the PP_{asym} case, just as in the TP_{asym} case.

7. Symmetric Firms with Step Marginal Cost Model

In preceding sections we have assumed that all suppliers have a common, constant marginal cost c for production up to capacity. Now we extend our analysis into the more realistic case in which firms have multi-step marginal cost functions. Rudkevich, et al (1998) derives SFE results for a multi-step marginal cost model with identical firms. Here we focus on the two-step case in order to illustrate how our approach can be applied to multi-step settings. We assume that each firm has some low cost generation facilities, with marginal cost c_1, and some higher cost generation facilities, with marginal cost c_2. The industry as a whole has K units of capacity. This total capacity consists of X low cost "base load" units (with marginal cost equal to c_1) and K-X high cost units (with marginal cost equal to c_2). We consider two cases in this setting. In the first case, off-peak load level is greater than the base load level, so that (given symmetric production levels) a firm always utilizes all of its low cost generation and some of its high cost generation facilities. In the second case, the off-peak load is less than the base load, and peak load is higher than the base load, so that a firm uses high cost generation facilities for part of the trading period, but not for all of the trading period. These cases are depicted on Figure Four for a player.

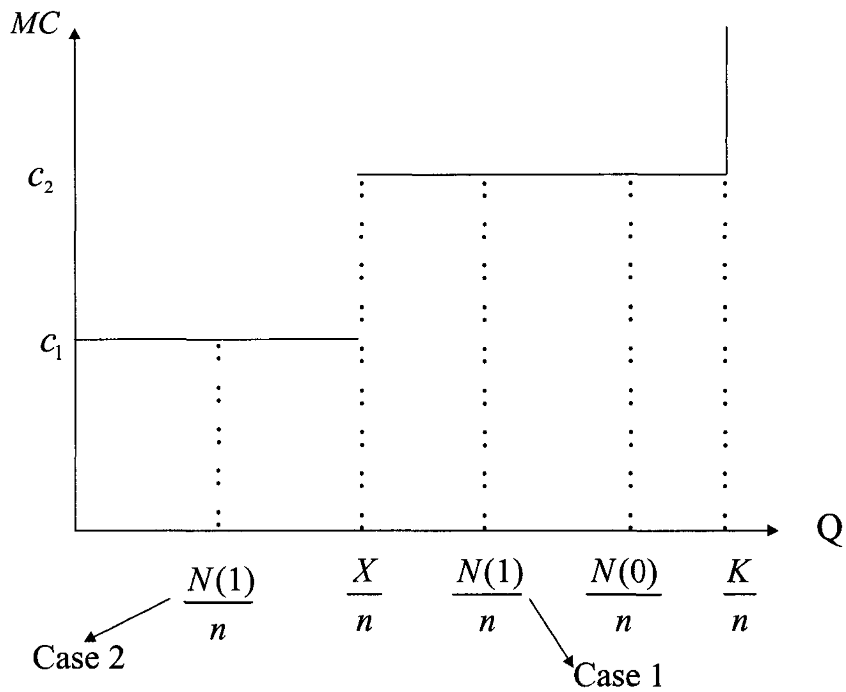
If we follow an analysis similar to that in section 4, we obtain the following differential equation, $S'(p) = \frac{S(p)}{(n-1)(p-z)}$, where $z(S(p))$ denotes the marginal cost function of a firm. This equality can be rewritten as $\frac{dp}{dQ} = \frac{(n-1)(p-z(Q))}{Q}$, where Q is the industry output. Solution of this ODE using initial price $p(Q^* = N(0)) = p_0$ satisfies $p(Q) = (Q)^{n-1}\left[\frac{p_0}{(Q^*)^{n-1}} + (n-1)\int_Q^{Q^*} \frac{z(x)}{x^n}dx\right]$, which is a special case of the price equation in Rudkevich et al (1998). As a function of time t, it can be rewritten as

(41) $$p(t) = (N(0)l^t)^{n-1}\left[\frac{p_0}{(N(0))^{n-1}} + (n-1)\int_{N(0)l^t}^{N(0)} \frac{z(x)}{x^n}dx\right].$$

Note that if $z(x) = c$, for all $x \geq 0$ then we obtain SFE equilibrium prices defined in section 4 as $p(t) = c + (p(0) - c)l^{t(n-1)}$.

Figure Four

Marginal Cost Function



Now for each case we examine how players' SFE and deviation profits vary. We characterize functions that determine the set of equilibria in the totally and partially pivotal regions with respect to base load capacity, capacity index, number of players, and load factor variations.

<u>Case 1</u> :

First we do the analysis for totally pivotal (TP) region. SFE profit in this case can be calculated as

(42)

$$\Pi^{SFE} = \int_0^1 \left[(p(t)-c_1)\frac{X}{n} + (p(t)-c_2)\left(\frac{N(t)-X}{n}\right)\right] dt = (c_2 - c_1)\frac{X}{n} + \frac{N(0)(p_0 - c_2)(l^n - 1)}{n^2 \ln l}.$$

The first integrand is the profit at time t meeting base load, and the second integrand is the profit for meeting the rest of the demand at time t. Note that this profit depends on the initial price $p(0)$ that indexes the candidate SFE, and it is a function of base load, cost parameters, load factor, number of players, but independent of the capacity index.

A deviating player's profit in the totally pivotal region (*TP*), which is defined in section 4, can be calculated as

(43)

$$\Pi^D = \int_t^1 \left(N(\tau) - \frac{(n-1)K}{n}\right)(\bar{p} - c_1) d\tau + \int_0^t \left[\left(N(\tau) - \frac{(n-1)K}{n}\right)(\bar{p} - c_2) + (c_2 - c_1)\frac{X}{n}\right] d\tau,$$

where t satisfies $t = \min\left\{1, \max\left\{0, \frac{\ln\left((x+k(n-1))/n\right)}{\ln l}\right\}\right\}$, where x is defined as

$x \equiv \frac{X}{N(0)}$, which is the ratio of the base load industry capacity, supplied by n firms, to the peak load, implying that $x < 1$. Larger values of x indicate that a higher portion of the demand is supplied at low cost. Also k is defined as $k \equiv \frac{K}{N(0)}$, as before, and $(k, l, n) \in TP$. This time t is determined as the time at which a deviating firm switches from high marginal cost production to low marginal cost production. If $t = 0$ then a deviating firm produces all of its output with low marginal cost. The first integral in Π^D calculates the deviating firm's profit for sales from time 1 to time t for which the firm is pivotal. The second integral, similarly, calculates the deviating firm's profit from sales for the rest of the time, given that t holds in the above equation. If we integrate the above expression we obtain

(44)

$$\Pi^D = N(0)\left(\bar{p}-c_1\right)\left[\frac{l-l^t}{\ln l}-(1-t)\frac{k(n-1)}{n}\right]+N(0)\left(\bar{p}-c_2\right)\left[\frac{l^t-1}{\ln l}-t\frac{k(n-1)}{n}\right]+t(c_2-c_1)\frac{X}{n},$$

with t as defined above.

Next to see how some of the candidate SFE equilibria may be ruled out we set the following relation:

$\Pi^{SFE} = \Pi^D$, or equivalently

(45) $\quad \gamma(.) \equiv (\bar{p}-c_1)\gamma_1(.)+(\bar{p}-c_2)\gamma_2(.)+(c_2-c_1)\gamma_3(.) = (p_0-c_2)$, where

$$\gamma_1(.) = \frac{n^2(l-l^t)-n(n-1)(1-t)k\ln l}{l^n-1},\ \gamma_2(.) = \frac{n^2(l^t-1)-n(n-1)tk\ln l}{l^n-1},\ \gamma_3(.) = \frac{xn(t-1)\ln l}{l^n-1}.$$

The function $\gamma(.)$ determines which intervals of markups are ruled out as equilibria for each vector of parameters, $(k,l,n,x,c_1,c_2,\bar{p})$. In particular, candidate SFE with markups of price over c_2 that are less than $\gamma(.)$ are ruled out as equilibria.

In the following example we show some cases in which for some parameters deviation is attractive, while for other parameters it is not.

Example 3: In this example we set up three problems with different set of parameters as in the table below. Here the triple (*k*, *l*, *n*) is in the totally pivotal region. In the first two problems (Pr. 1, Pr.2), we begin with a larger base load share (which means that most of the time less costly generators will be utilized) and compare the effects of

Table One: Factor function values that determine whether candidate equilibria ruled out or not in the totally pivotal region.

	k	**l**	**n**	**x**	$\bar{p}$	c_1	c_2	γ_1	γ_2	γ_3	γ
Pr. 1	1.15	0.8	3	0.6	100	30	90	.398	.135	-.698	-12.6
Pr. 2	1.15	0.8	3	0.6	100	5	10	.398	.135	-.698	46.5
Pr. 3	1.15	0.77	3	0.4	100	15	80	.172	.318	-.345	-1.39

$(\bar{p}-c_1)$, and $(\bar{p}-c_2)$ on the deviation incentives. A negative γ value means that no candidate SFE will be ruled out. By the same token, a positive value of γ implies that deviation is more profitable than bidding supply functions. Particularly, Pr.2 shows that candidate SFE with initial prices less than 56.5 are ruled out. In problem 3 (Pr. 3), we decrease the base load factor and change the marginal cost parameters. The negative value of γ for Pr.3 indicates that no SFE with a initial price $p_0 \in [c_2, \bar{p}] = [80,100]$ can be ruled out by a simple deviation.

Next we perform the analysis for partially pivotal (PP) region in Case 1.In this case the partially pivotal region is as defined in section 4. A deviating player's profit can be calculated as

(46)

$$\hat{\Pi}^D = \int_t^\tau \left(N(T) - \frac{(n-1)K}{n} \right) (\bar{p} - c_1) dT + \int_0^t \left[\left(N(T) - \frac{(n-1)K}{n} \right) (\bar{p} - c_2) + (c_2 - c_1) \frac{X}{n} \right] dT ,$$

where t satisfies $t = \min\left\{ \tau, \max\left\{ 0, \frac{\ln\left((x + k(n-1))/n\right)}{\ln l} \right\} \right\}$, and τ satisfies

$\tau = \frac{\ln\left((n-1)k/n\right)}{\ln l}$. The player is pivotal between 0 and τ, and residual demand is the one as specified above for any time less than τ, and residual demand is zero for any time value above τ. Note that $t, \tau \in [0,1]$ and $\tau \geq t$. Interpretation of the integrands is similar to the ones defined above. We take the integral and end up with

(47)

$$\hat{\Pi}^D = N(0)(\bar{p} - c_1) \left[\frac{l^\tau - l^t}{\ln l} - (\tau - t) \frac{k(n-1)}{n} \right] + N(0)(\bar{p} - c_2) \left[\frac{l^t - 1}{\ln l} - t \frac{k(n-1)}{n} \right] + t(c_2 - c_1) \frac{X}{n} ,$$

where the triple $(k, l, n) \in PP$.

Now we show how some of the candidate SFE may be ruled out in the partially pivotal region. We set the relation $\hat{\Pi}^D = \Pi^{SFE}$, or equivalently

(48) $\quad e(.) \equiv (\bar{p} - c_1) e_1(.) + (\bar{p} - c_2) e_2(.) + (c_2 - c_1) e_3(.) = (p_0 - c_2)$, where

$$e_1(.) = \frac{n^2(l^\tau - l^t) - n(n-1)(\tau - t)k\ln l}{l^n - 1},\ e_2(.) = \frac{n^2(l^t - 1) - n(n-1)tk\ln l}{l^n - 1},\ e_3(.) = \frac{xn(t-1)\ln l}{l^n - 1}.$$

Again $e(.)$ produces a function that determines which interval of markups is ruled out as equilibria for each vector of parameters, $(k, l, n, x, c_1, c_2, \overline{p})$.

Proposition 6: If $X < N(1)$ then as the base load capacity, X, increases, deviation becomes less attractive, or as the base load capacity decreases, a pivotal firm's deviation incentive increases. That is $\partial\gamma(.)/\partial x < 0$ for $(k, l, n) \in TP$, and $\partial e(.)/\partial x < 0$ for $(k, l, n) \in PP$.

Proof – see appendix.

An interpretation of this proposition is that as the base load factor, *x*, increases profit from SFE increases, because each firm produces more at the low cost level. Also as *x* increases time *t* decreases. As *t* goes down a deviating player's profit also increases, because more of the time he meets the residual demand at the low cost. However this rate of increase is less than that of increase in profit from SFE. Hence a pivotal firm's deviation incentive decreases.

Case 2 :

We first analyze this case in the totally pivotal (TP) region. In this case one can compute SFE profit as follows:

$$\text{(49)} \qquad \Pi^{SFE} = \int_{t^*}^{1} (p(t) - c_1)\frac{N(t)}{n}dt + \int_{0}^{t^*} \left[(p(t) - c_1)\frac{X}{n} + (p(t) - c_2)\left(\frac{N(t) - X}{n}\right) \right] dt,$$

where t^* satisfies $N(t^*) = X$, that is $t^* = \ln x / \ln l$, and equilibrium prices are defined as

(50) $$p(t)=\begin{cases} c_2+(p_0-c_2)l^{t(n-1)} & , \; if \; t<t^* \\ c_1+\left[p_0-c_2+x(c_2-c_1)\right]l^{t(n-1)} & , \; if \; t>t^*. \end{cases}$$

These prices are obtained from equation (41). One can easily interpret the integrands as above.

A deviating player's profit in the totally pivotal region can also be obtained as

(51)

$$\Pi^D=\int_{t'}^{1}\left(N(t)-\frac{(n-1)K}{n}\right)(\bar{p}-c_1)dt+\int_{0}^{t'}\left[(\bar{p}-c_1)\frac{X}{n}+\left(N(t)-\frac{X}{n}-\frac{(n-1)K}{n}\right)(\bar{p}-c_2)\right]dt,$$

where t' holds $t'=\max\left\{t^*,\max\left\{0,\frac{\ln\left((x+(n-1)k)/n\right)}{\ln l}\right\}\right\}$. Note that $t^*\le t'$ and $l<x$ always hold for $x>0$ in the TP region.

Now we can define the equation that determines which interval of markups is ruled out as equilibria. We first integrate the above expressions;

(52) $$\Pi^{SFE}=\frac{(p_0-c_2)N(0)}{n^2\ln l}\left[l^n-1\right]+\frac{(c_2-c_1)X}{n^2\ln l}\left[t^*n\ln l+l^n-l^{nt^*}\right],$$ and

(53)

$$\Pi^D=N(0)(\bar{p}-c_1)\left[\frac{l-l^{t'}}{\ln l}-(1-t')\frac{k(n-1)}{n}\right]+N(0)(\bar{p}-c_2)\left[\frac{l^{t'}-1}{\ln l}-t'\frac{k(n-1)}{n}\right]+t'(c_2-c_1)\frac{X}{n}$$

then we set the relation $\Pi^{SFE}=\Pi^D$, which implies that

(54) $$o(.)\equiv(\bar{p}-c_1)o_1(.)+(\bar{p}-c_2)o_2(.)+(c_2-c_1)o_3(.)=(p_0-c_2),$$ where

$$o_1(.) = \frac{n^2(l-l^{t'}) - n(n-1)(1-t')k\ln l}{l^n - 1}, \quad o_2(.) = \frac{n^2(l^{t'}-1) - n(n-1)t'k\ln l}{l^n - 1},$$

$$o_3(.) = \frac{x\left((t'-t^*)n\ln l - l^n + l^{nt^*}\right)}{l^n - 1}.$$

Now we will investigate the relationship between the function $o(.)$ and base load factor x, as we performed in proposition 6 for Case 1. However the below example shows that the rate of change of the function with respect to the base load factor can go either direction depending on the values of the parameters.

<u>*Example 4*</u> : First we derive the equation for $\partial o/\partial x$. After some algebra one can obtain that $\dfrac{\partial o}{\partial x} = \dfrac{(c_2 - c_1)}{l^n - 1}\left[\ln\left(\dfrac{x + k(n-1)}{nx}\right)^n + n(x^n - 1) + x^n - l^n\right]$.

Now the examples:

i) For $n = 2$, $k = 1$, $l = 0.5$, the derivative at $x = 0.9$ becomes $\dfrac{\partial o}{\partial x} = -0.38(c_2 - c_1)$.

ii) For $n = 2$, $k = 1$, $l = 0.55$, the derivative at $x = 0.6$ becomes $\dfrac{\partial o}{\partial x} = 0.92(c_2 - c_1)$.

Hence, from this example it is clear that deviation becomes more or less attractive as base load capacity changes depending on the values of the vector (k, l, n, x) given that marginal cost parameters are fixed.

We have also searched for the properties of the function that determines whether the deviation is attractive or not in the partially pivotal region for Case 2. We have obtained a similar observation that we had for the totally pivotal region. That is the rate of change of the function with respect to the base load factor can go either direction depending on the values of the parameters. For the sake of brevity we have omitted the derivations and examples for Case 2.

8. Conclusions

SFE is a widely used tool in the electricity industry to study generators' bidding behaviors and market power issues. Effectiveness of pivotal suppliers on the market outcomes is also known in the industry as well. However SFE with pivotal suppliers was not considered in the literature.

In this paper we examined the connection between pivotal suppliers and the set of SFE. In the constant marginal cost, symmetric firms model, we showed that when pivotal suppliers are present, the set of SFE would shrink. The most competitive equilibria are eliminated as the number of players and the capacity index decrease and the load factor increases. In the step marginal cost symmetric model, depending on the position of the base load, we characterized the SFE with respect to the change in the base load factor. In particular, we showed that deviation from candidate SFE becomes more (less) attractive; as the base load decreases (increases) in which base load is less than the off-peak load level. This holds whether pivotal players are in the totally or partially pivotal regions.

We also discussed optimal deviations in the symmetric firms model. We illustrated a case in which a simple kind of deviation discussed in this paper cannot rule out some SFE, which would be ruled out if an optimal deviation is implemented. In the asymmetric firms model we showed that the larger firm has an incentive to deviate from a wider range of SFE, and it is the larger firm's deviation incentives that determine which candidate SFE are ruled out as equilibrium. We also showed that more (less) candidate SFE is ruled out as load factor increases (decreases) and capacity index decreases (increases) in the totally and partially pivotal regions.

As a future research we plan to test the findings of this paper in the England and Wales electricity market.

APPENDIX B

Proof of Proposition 1: We begin by listing several facts about functions of parameters (l, n) that we use in the proof. We omit the proof of a fact when the proof is straightforward.

Fact 1: $1 - l^n + nl^n \ln l > 0$ for $l < 1$.

Fact 2: $1 - l + \ln l < 0$ for $l < 1$.

Fact 3: $b(l,n) \equiv n(1-l) + n\ln l - \ln l \in (0,(1-l)/l)$ for $l \in ((n-1)/n,1)$.

Proof of Fact 3: $b(1,n) = 0$ and $\partial b(l,n)/\partial l = -n + (n-1)/l < 0$ since $l \in ((n-1)/n,1)$. As a result, $b(l,n) > 0$. Next we consider the proposed upper bound for $b(l,n)$, $a(l) \equiv (1-l)/l$. We note that $a(1) \equiv 0$. One can show that $a'(l) < \partial b(l,n)/\partial l < 0$ for $l \in ((n-1)/n,1)$. This slope condition, coupled with the condition $b(1,n) = a(1) = 0$, establishes that $b(l,n) < a(l)$ for $l \in ((n-1)/n,1)$.

Fact 4: $(n-1)\ln l + 1 \in (0,l)$ for $l \in ((n-1)/n,1)$.

Proof of Fact 4: First note that $e^{-1/(n-1)} < \dfrac{n-1}{n}$ for $n \geq 2$.Taking logs of both sides of the inequality yields, $-1/(n-1) < \ln((n-1)/n) < \ln l$, where the second inequality follows from the lower bound on l. Rearranging this inequality yields $(n-1)\ln l + 1 > 0$. Next we consider the upper bound for $(n-1)\ln l + 1$. Note that $\ln l < l - 1$ for $l < 1$. Furthermore, $(n-1)\ln l \leq \ln l$ for $l < 1$, since $n \geq 2$. Combining these two inequalities yields, $(n-1)\ln l < l - 1$, or $(n-1)\ln l + 1 < l$, which is the desired result.

The remainder of the proof is divided into three parts, with one part devoted to the impact of each element of (k,l,n).

(*a*) *Capacity index k*: Note that $\phi(.)$ is an affine function in *k*, and

$\partial\phi/\partial k = [n(n-1)\ln l]/[1-l^n] < 0$, since $\ln l < 0$.□

(*b*) *Load factor l*: We initially suppose that $k = 1$. At $k = 1$ we have,

$$\text{(A1)} \quad l(1-l^n)^2 \frac{\partial\phi}{\partial l} = l^n[(n^3-n^2)\ln l + l(n^2-n^3) + n^3 - n^2 + n] + [n^2(1-l) - n].$$

Define the function $\psi(l,n)$ as the right-hand-side of equation (A1). Note that $\psi(1,n) = 0$. The partial of ψ with respect to the load factor simplifies to,

$\partial\psi/\partial l = n^2 l^n[(n(n-1)\ln l)/l - n^2 + 1 + n^2/l] - n^2$.

Let $q(l) = [n(n-1)\ln l/l - n^2 + 1 + n^2/l] - l^{-n}$. We claim that $q(l) < 0$. But this holds since $q(l=1) = 0$, and, $\partial q/\partial l = -n[(n-1)\ln l + 1]/l^2 + n[l^{-n+1}]/l^2 > 0$ because of fact 4. This in turn implies that $\psi_l(l,n) < 0$ for $l < 1$, so that $\psi(l,n) > 0$ for $l < 1$. Using (**), we have $\phi_l(1,l,n) > 0$.

Now we extend this result to $k > 1$. Differentiating ϕ_l with respect to *k* yields,

$$\phi_{lk} = \frac{1}{(1-l^n)^2}\left((1-l^n)\left(\frac{n(n-1)}{l}\right) + n(n-1)nl^{n-1}\ln(l)\right) = \frac{n(n-1)}{l(1-l^n)^2}\left[1 - l^n + nl^n \ln(l)\right]$$

The term in square brackets is positive by Fact 1, so that, $\phi_{lk} > 0$. This result, coupled with $\phi_l(1,l,n) > 0$, implies that $\phi_l(k,l,n) > 0$ for all $(k,l,n) \in TP$.□

(c) *Number of suppliers n*: We initially suppose that $k = 1$. At $k = 1$ we have,

$$\text{(A2)} \quad (1-l^n)^2 \frac{\partial \phi}{\partial n} = \{1-l^n + nl^n \ln l\}[n(1-l) + n\ln l - \ln l] + n(1-l^n)[1-l+\ln l] \; .$$

The sign of $\partial\phi/\partial n$ is not clear, because the first term on the RHS of (A2) is positive (by Facts 1 and 3) and the last term is negative (by Fact 2). Define the function $\omega(l,n)$ as the right-hand-side of equation (A2). Note that $\omega(1,n)=0$. The partial of ω with respect to the load factor may be expressed as,

$$\partial\omega(l,n)/\partial l = E+F+G+H \, ,$$

where, $E=(n^2 l^{n-1}\ln l)(n(n-1)+(n-1)\ln l)$, $F=(1-l^n+nl^n\ln l)(-n+(n-1)/l)$, $G=(-n^2 l^{n-1})(1-l+\ln l)$, and $H=n(1-l^n)(-1+1/l)$.

By using the facts stated at the beginning of the proof we can establish the following inequalities: $E<0$, $F>0$, $G>0$, and $H>0$. Let $D \equiv E+H$. Note that $D=0$ at $l=1$. Consider the derivative of D with respect to the load factor,

$$\partial D/\partial l = \partial E/\partial l + \partial H/\partial l \, ,$$

where,

$$\partial E/\partial l = (n^2 l^{n-2}(1+(n-1)\ln l))(n(1-l)+(n-1)\ln l) + (n^2 l^{n-1}\ln l)(-n+(n-1)l^{-1}) > 0 \, ,$$

since the first term on the RHS is the product of two positive terms (using Facts 3 and 4) and the second term is the product of two negative terms, and

$$\partial H/\partial l = (-n^2 l^{n-1}(-1+1/l)) + (n-nl^n)(-1/l^2) < 0 \, .$$

Next we compare the magnitude of $\partial E/\partial l$ and $\partial H/\partial l$. Comparing the first terms in these two derivatives, we have,

$$\text{(A3)} \quad \left|-n^2 l^{n-1}(-1+1/l)\right| > (n^2 l^{n-2}(1+(n-1)\ln l))(n(1-l)+(n-1)\ln l) \, ,$$

by using Facts 3 and 4. Define the function $m(l,n)$ as the sum of the last term in $\partial E/\partial l$ and the last term in $\partial H/\partial l$, so that,

$$m(l,n) = (n^2 l^{n-1} \ln l)(-n + (n-1)l^{-1}) + (n - nl^n)(-1/l^2).$$

By using $m(1,n) = 0$ combined with the kind of derivative argument that we have used several times in this proof, we can establish that $m(l,n) < 0$. This fact, coupled with the inequality (A3), yields, $\partial D/\partial l = \partial E/\partial l + \partial H/\partial l < 0$, which in turn establishes that $D \equiv E + H > 0$. Signing D was the final step in establishing that $\partial\omega(l,n)/\partial l = E + F + G + H > 0$. This derivative result for $\omega(l,n)$ establishes that $\omega(l,n) < 0$. Since we defined the LHS of (A2) as $\omega(l,n)$, we have shown that $\partial\phi/\partial n$ is negative, evaluated at $k = 1$.

The final step in the proof of part (c) is to extend the derivative result for $\partial\phi/\partial n$ to $k > 1$. We have,

$$(1-l^n)^2 \frac{\partial\phi}{\partial n} = (1 - l^n + nl^n \ln l)[n(1-l) + kn\ln l - k\ln l] + n(1-l^n)[1 - l + k\ln l],$$

which is affine in k. Also,

$$(1-l^n)^2 \frac{\partial^2\phi}{\partial n\partial k} = [1 - l^n + nl^n \ln l](n-1)\ln l + (n - nl^n)\ln l < 0 \;,$$

where the inequality follows because the term in square brackets is positive by Fact 1. As k increases $\frac{\partial\phi}{\partial n}$ becomes more negative, thus the result of part (c) follows. □

Proof of Proposition 2:

Let $h(k,l,n,\tau) \equiv [n^2(1-l^\tau)+n(n-1)k\tau \ln l]/[1-l^n]$. Then

$\partial h/\partial \tau = n \ln l[(n-1)k - nl^\tau]/(1-l^n) = 0$, since the term in square brackets is zero by the definition of τ. We will use this equality for the proof of the following.

(a) *Capacity index k:* We take the total derivative of $\lambda(.)$ with respect to *k*.

$$\frac{d\lambda}{dk} = \frac{\partial h}{\partial \tau}\frac{\partial \tau}{\partial k} + \frac{\partial \lambda}{\partial k}.$$

Here $\partial\lambda/\partial k = n(n-1)\tau \ln l/(1-l^n) < 0$, since $\ln l < 0$. Thus $\dfrac{d\lambda}{dk} < 0$.□

(b) *Load factor l:* We take the total derivative of $\lambda(.)$ with respect to *l*.

We obtain $\dfrac{d\lambda}{dl} = \dfrac{\partial h}{\partial \tau}\dfrac{\partial \tau}{\partial l} + \dfrac{\partial \lambda}{\partial l}$, where we need to determine the sign of second term $\partial\lambda/\partial l$.

We have

$$l(1-l^n)^2\frac{\partial \lambda}{\partial l} = l^n[\tau k(n^3-n^2)\ln l + l^\tau(\tau n^2 - n^3) + n^3 - \tau k n^2 + \tau k n] + [\tau n^2(k-l^\tau) - \tau k n].$$

Note that $l(1-l^n)^2 \left.\dfrac{\partial\lambda}{\partial l}\right|_{\tau=0} = 0$.

(B1) The last term, $\tau n^2(k-l^\tau) - \tau k n = \tau n^2[k(n-1)/n - l^\tau]$ is equal to zero, by the definition of τ.

(B2) Let $Y(.) = [\tau k(n^3-n^2)\ln l + l^\tau(\tau n^2 - n^3) + n^3 - \tau k n^2 + \tau k n]$. If we show that $\partial Y/\partial \tau < 0$, proof completes. We obtain

$\partial Y/\partial \tau = k(n^3-n^2)\ln l + nk(1-n) + l^\tau n^2 + l^\tau(\tau n^2 - n^3)\ln l$. Denote $y = \partial Y/\partial \tau$, then $\partial y/\partial k = (n^3-n^2)\ln l + n(1-n) < 0$, implies that *y* reaches its maximum at *k*=1. Now re-write *y* at *k*=1 as $\bar{y} = (n^3-n^2)\ln l + n(1-n) + l^\tau n^2 + l^\tau(\tau n^2 - n^3)\ln l$. We will show that maximum of $\bar{y}$ is negative. Then $\partial\bar{y}/\partial\tau = l^{\tau}n^{2}\ln l[2+(\tau-n)\ln l]<0$. Now if we solve $\bar{y}=0$ for τ, we obtain $\bar{\tau}=\ln(1+1/[(n^{2}-n)\ln l-n])/\ln l$, but $\bar{\tau}$ is less than the constraint $\tau(k=1)=\ln((n-1)/n)/\ln l$. These imply that $\partial Y/\partial\tau<0$.

Thus by the results of *(B1)* and *(B2)*, the proof of this proposition is complete. □

(c) *Number of suppliers n:* We take the total derivative of $\lambda(.)$ with respect to n.

We obtain $\dfrac{d\lambda}{dn}=\dfrac{\partial h}{\partial\tau}\dfrac{\partial\tau}{\partial n}+\dfrac{\partial\lambda}{\partial n}$, where we need to determine the sign of second term. Now if we show that for all τ, the maximum of $\partial\lambda/\partial n$ is negative then the proof is complete.

The partial derivative of λ with respect to n becomes

$(1-l^{n})^{2}\dfrac{\partial\lambda}{\partial n}=(1-l^{n}+nl^{n}\ln l)[n(1-l^{\tau})+\tau k(n-1)\ln l]+n(1-l^{n})[1-l^{\tau}+\tau k\ln l]$. Note that

$(1-l^{n})^{2}\left.\dfrac{\partial\lambda}{\partial n}\right|_{\tau=0}=0$. Now consider how the partial derivative of λ with respect to n varies

with τ: $(1-l^{n})^{2}\dfrac{\partial^{2}\lambda}{\partial n\partial\tau}=(1-l^{n}+nl^{n}\ln l)[-nl^{\tau}\ln l+k(n-1)\ln l]+n(1-l^{n})(-l^{\tau}\ln l+k\ln l)$.

The term in square brackets is zero, given the definition of τ. The fact that $k\geq 1>l^{\tau}$ implies that $(1-l^{n})^{2}\dfrac{\partial^{2}\lambda}{\partial n\partial\tau}<0$. Since $\left.\dfrac{\partial\lambda}{\partial n}\right|_{\tau=0}=0$ and $\dfrac{\partial^{2}\lambda}{\partial n\partial\tau}<0$, we have $\dfrac{\partial\lambda}{\partial n}<0$ for $\tau=\ln[k(n-1)/n]/\ln l>0$. As a consequence, the total derivative of $\lambda(.)$ with respect to n becomes negative. □

Proof of Proposition 3:

Let $v = \psi_1 - \psi_2 = \frac{4(l-1-k_2 \ln l)}{l^2 - 1 + 2\ln l - 4k_2 \ln l} - \frac{4(l-1-k_1 \ln l)}{l^2 - 1 - 2\ln l + 4k_2 \ln l}$. We will show that even minimum of v is positive.

$\frac{\partial \psi_1}{\partial k_2} < 0$, since $sign\left(\frac{\partial \psi_1}{\partial k_2}\right) = sign\left((-4\ln l)(1-\psi_1)(l^2 - 1 + 2\ln l - 4k_2 \ln l)\right) < 0$, also

$\frac{\partial \psi_2}{\partial k_2} = \frac{4(l-1-k_1 \ln l)4\ln l}{\left(l^2 - 1 - 2\ln l + 4k_2 \ln l\right)^2} > 0$. Thus $\frac{\partial v}{\partial k_2} < 0$.

$\frac{\partial v}{\partial k_1} = \frac{4\ln l}{\left(l^2 - 1 - 2\ln l + 4k_2 \ln l\right)} > 0$. Hence minimum of v is attained at $k_1 = 0.5 = k_2$,

moreover $v(k_1 = 0.5 = k_2) = 0$. But $v(.)\big|_{k_2 = 0.5} = \frac{4(k_1 - k_2)\ln l}{l^2 - 1} > 0$, since $k_1 > k_2$. Hence the result. □

Proof of Proposition 4:

(a) *Capacity index* k_2: We take the first derivative of $\psi_1(.)$ with respect to k_2. After suitable arrangements and substitutions, we end up with

$sign\left(\frac{\partial \psi_1}{\partial k_2}\right) = sign\left((-4\ln l)(1-\psi_1)(l^2 - 1 + 2\ln l - 4k_2 \ln l)\right) < 0$, since the first two terms are positive, and the final term is negative. □

(b) *Load factor* l: After taking first derivative of $\psi_1(.)$ with respect to l, we obtain that

$$sign\left(\frac{\partial \psi_1}{\partial l}\right) = sign\left(4(1 - k_2 / l)(l^2 - 1 + 2\ln l - 4k_2 \ln l) - 4(l - 1 - k_2 \ln l)(2l + 2/l - 4k_2 / l)\right)$$

The sign is not clear yet, because first and last terms are positive and others are negative.

Let $Z(.) = 4(1-k_2/l)(l^2-1+2\ln l-4k_2\ln l)-4(l-1-k_2\ln l)(2l+2/l-4k_2/l)$. We will show that minimum of Z(.) is positive. Note that

$Z(l=k_2)>0$, and moreover

$\frac{\partial Z}{\partial l} = 4(-k_2/l^2)(l^2-1+2\ln l-4k_2\ln l)-4(l-1-k_2\ln l)(2-2/l^2+4k_2/l^2)>0$, since the first three terms in the parenthesis are negative, and the last term in the parenthesis is positive. Hence the result follows. □

Proof of Proposition 5:

Let $b(k_2,l,\hat{\tau}) \equiv \frac{4(l^{\hat{\tau}}-1-\hat{\tau}k_2\ln l)}{(l^2-1+2\ln l-4k_2\ln l)}$. Then $\frac{\partial b}{\partial \hat{\tau}} = \frac{4\ln l\left(l^{\hat{\tau}}-k_2\right)}{\left(l^2-1+2(1-2k_2)\ln l\right)}=0$, because numerator is zero by the definition of $\hat{\tau}$, and denominator is negative. We use this relation for the proof of the following.

(a) *Capacity index* k_2: If we take the total derivative of $\beta_1(.)$ with respect to k_2, we obtain $\frac{d\beta_1}{dk_2} = \frac{\partial b}{\partial \hat{\tau}}\frac{\partial \hat{\tau}}{\partial k_2}+\frac{\partial \beta_1}{\partial k_2}$, where

$sign\left(\frac{\partial \beta_1}{\partial k_2}\right) = sign\left(-4(\hat{\tau}\ln l)(l^2-1+2(1-2k_2)\ln l)+4(l^{\hat{\tau}}-1-\hat{\tau}k_2\ln l)(4\ln l)\right)$. Here all terms in parenthesis are negative, hence overall sign is not clear yet.

Let $B(.) = 4(-\hat{\tau}\ln l)(l^2-1+2(1-2k_2)\ln l)-4(l^{\hat{\tau}}-1-\hat{\tau}k_2\ln l)(-4\ln l)$. We make suitable substitutions into B(.), take the partial derivative of it with respect to l, and after rearrangements, we end up with $sign\left(\frac{\partial B}{\partial l}\right) = sign\left(-8l^2 \ln k_2 - 8\ln k_2 + 16k_2 + 16\right)$, which is obviously positive. Also note that $B(l = k_2) = -4\ln k_2\left(k_2^2 - 4k_2 + 3 + 2\ln k_2\right)$, in which the sign is not obvious. Let $bb(k_2) = k_2^2 - 4k_2 + 3 + 2\ln k_2$, where $bb' > 0$, and $bb(k_2 = 0.5) < 0$. Thus $B(l = k_2) < 0$. These imply that the β function is decreasing in the (firm 2's) capacity index k_2.□

(b) *Load factor l*: Total derivative of $\beta_1(.)$ with respect to l becomes

$\frac{d\beta_1}{dl} = \frac{\partial b}{\partial \hat{\tau}}\frac{\partial \hat{\tau}}{\partial l} + \frac{\partial \beta_1}{\partial l}$, where

$$sign\left(\frac{\partial \beta_1}{\partial l}\right) = sign\left(4(\hat{\tau}l^{\hat{\tau}-1} - \hat{\tau}k_2 / l)(l^2 - 1 + 2(1-2k_2)\ln l) - 4(l^{\hat{\tau}} - 1 - \hat{\tau}k_2 \ln l)(2l + 2(1-2k_2)/l)\right)$$

which is positive, since first term in parenthesis is zero by definition of $\hat{\tau}$, and the last two terms are positive.□

Proof of Corollary 5:

The $B(.)$ function defined in the proof of proposition 5 can be rewritten as

$B(.) = \left(\frac{\hat{\tau}}{\beta} - 1\right)4(l^{\hat{\tau}} - 1 - \hat{\tau}k_2 \ln l)(-4\ln l)$. Multiplication of last two terms is negative, and we have proven above that $B(.)$ is negative, these imply that first term on the above equation must be positive, hence the result. □

Proof of Proposition 6:

i) $(k, l, n) \in TP$:

For $t > 0$, $\frac{\partial t}{\partial x} = \frac{1}{\left(x+(n-1)k\right)\ln l} < 0$, since $\ln l$ is negative. Then

$\frac{\partial \gamma_3}{\partial x} = \frac{n \ln l}{l^n - 1}\left((t-1) + x\frac{\partial t}{\partial x}\right) < 0$ holds. One can obtain

$\frac{\partial \gamma_2}{\partial x} = \frac{\partial t}{\partial x} \ln l \left[n^2 l^t - n(n-1)k\right] \frac{1}{l^n - 1} < 0$, because the term in square brackets is positive.

Also note that $\frac{\partial \gamma_1}{\partial x} = -\frac{\partial \gamma_2}{\partial x}$. Then

$\frac{\partial \gamma}{\partial x} = (\overline{p} - c_1)\frac{\partial \gamma_1}{\partial x} + (\overline{p} - c_2)\frac{\partial \gamma_2}{\partial x} + (c_2 - c_1)\frac{\partial \gamma_3}{\partial x} = (c_2 - c_1)\left[\frac{\partial \gamma_1}{\partial x} + \frac{\partial \gamma_3}{\partial x}\right]$ holds. The sign of this term is not clear yet, because in the square brackets one term is positive and the other term is negative. After suitable substitutions, and rearrangements we obtain that

$\frac{\partial \gamma_1}{\partial x} = \frac{n \ln l}{l^n - 1}\frac{\partial t}{\partial x}\left[-x\right]$, and $\frac{\partial \gamma_3}{\partial x} = \frac{n \ln l}{l^n - 1}\frac{\partial t}{\partial x}\left[\frac{t-1}{\partial t / \partial x} + x\right]$. We add them up,

$\frac{\partial \gamma_1}{\partial x} + \frac{\partial \gamma_3}{\partial x} = \frac{n \ln l}{l^n - 1}\frac{\partial t}{\partial x}\left[\frac{t-1}{\partial t / \partial x}\right] < 0$, since the first and last terms are positive, and the second term is negative. Hence the result follows.

If $t = 0$ then $\frac{\partial t}{\partial x} = 0$, and $\frac{\partial \gamma}{\partial x} = \frac{\partial \gamma_3}{\partial x} = \frac{-n \ln l}{l^n - 1} < 0$.

ii) $(k, l, n) \in PP$:

Proof is very similar to the proof of part i), because $\frac{\partial \tau}{\partial x} = 0$, and $\tau > t$. For $t > 0$, again

$\frac{\partial e_3}{\partial x} < 0$, $\frac{\partial e_2}{\partial x} < 0$, and $\frac{\partial e_1}{\partial x} = -\frac{\partial e_2}{\partial x}$, because of the similar reasons used above proof.

Finally one can obtain $\frac{\partial e}{\partial x} = (c_2 - c_1)\left[\frac{\partial e_1}{\partial x} + \frac{\partial e_3}{\partial x}\right]$, where

$\frac{\partial e_1}{\partial x} + \frac{\partial e_3}{\partial x} = \frac{n \ln l}{l^n - 1} \frac{\partial t}{\partial x}\left[\frac{t-1}{\partial t / \partial x}\right] < 0$, hence the result. For other t values the result of the

proof remains unchanged □

REFERENCES

Arena. http://www.software.rockwell.com

Allaz, B., J.L. Vila, "Cournot Competition, Futures Markets and Efficiency", Journal of Economic Theory, vol.59, 1993.

Baldick, R. and W. Hogan, "Capacity Constrained Supply Function Equilibrium Models of Electricity Markets: Stability, Non-decreasing Constraints, and Function Space Iterations", PWP-089 Working Paper, August 2002.

Beale, E.M.L., "On minimizing a convex function subject to linear inequalities", *J. Royal Statistical Society*, Series B, v. 17, 1955.

Bellman, R., "Some applications of the theory of dynamic programming - a review" *Operations Research*, v. 2, 275-288, 1954.

Berry, S. and A. Pakes, " Some Applications and Limitations of Recent Advances in Empirical Industrial Organization: Merger Analysis" *American Economic Review*, v. 83, May 1993.

Bertsekas, D. P., *Dynamic Programming,* Prentice Hall, 1987.

Birge, J.R and F. Louveaux, *Introduction to Stochastic Programming,* Springer, 1997.

Blackburn, T., "New Rules for Market-Priced Sales Transactions", Mimeo, Bruder-Gentle & Marcoux L.L.P., 2001.

Bollerslev, T. "Generalized Autoregressive Conditional Heteroskedasticity", *Journal of Econometrics,* vol.31, pp. 307-327, 1986.

Borenstein, S., J. Bushnell, E. Kahn, S. Stoft, "Market Power in California Electric Markets", *Utilities Policy*, vol.5, 1996.

Borenstein, S., J. Bushnell, "An Empirical Analysis of the Potential for Market Power in the California's Electricity Industry", NBER Working Paper 6463, 1998.

Borenstein, S., Bushnell, J., "In the Dark California Consumers Haven't Seen Benefits of Deregulating the Electrical Industry Yet—What Went Wrong?" , *Los Angeles Times*, August 27, 2000.

Borenstein, S., Bushnell, J.(2001) "Electricity Restructuring: Deregulation or Regulation?", *Regulation*, volume 23. No.2, pp. 46-52.

Borenstein, S., "The Trouble With Electricity Markets and California's Electricity Restructuring Disaster", PWP-081, September 2001.

Bushnell, J., C.R. Knittel, F. Wolak, "Estimating the Opportunities for Market Power in a Deregulated Wisconsin", Mimeo, University of California Energy Institute, 1998.

Cardell, J.B., C.C. Hitt, W.W. Hogan, "Market Power and Strategic Interaction in Electricity Networks", *Resource and Energy Economics*, vol. 19, pp. 109-137, 1997.

Carino, D. and W. Ziemba, "Formulation of Russell-Jasuda Karsai Financial Planning Model" *Operations Research*, v. 46, No 4, 1998.

Casey, M. and S. Sen, "The Scenario Generation Algorithm for Multi-stage Stochastic Linear Programming" manuscript, University of Arizona, 2002. (http://www.sie.arizona.edu/SPEED-CS/raptormore/more/papers.html)

Charnes, A. and W.W. Cooper, "Change-constrained programming" *Management Science*, v. 6:73, 1959.

Dantzig, G.B., "Linear programming under uncertainty" *Management Science*, v. 1, 1955.

Deneckere, R. J. and A. de Palma, "The diffusion of consumer durables in a vertically differentiated oligopoly" *RAND Journal of Economics,* v. 29, 1998

Engle, R. "Autoregressive Conditional Heteroskedasticity With Estimates of the Variance of U.K. Inflation", *Econometrica* 50, pp. 987-1008, 1982.

Ericson, R. and A. Pakes, "Markov Perfect Industry Dynamics: A framework for Empirical Work" *Review of Economic Studies*, v. 62, 1995.

Ewing K.P., Serota, J.I. "Antitrust and the continuing deregulation of the US energy industry", *The Antitrust Review of The Americas 2000: A Global Competition Review Special Report*, pp57-60, 2000.

Eydeland A., Geman H. "Some Fundamentals of Electricity Derivatives". Working paper, Southern Company Energy Marketing, and University Paris IX Dauphine and ESSEC, 1998, July.

Feinberg, E. "Electricity Power Load Pocket Modeling and Forecasting," University of Florida Seminar, March, 2002.

Fisher, M.J., J. Hammond, W. Obermeyer, and A. Raman, "Configuring a supply chain to reduce the cost of demand uncertainty" *Production and Operations Management*, v. 6, 1997.

Friedman, J.M., *Oligopoly Theory,* North Holland, 1977.

Forgo, F. and F. Szidarovsky, *Introduction to the Theory of Games*, Kluwer Academic Publishers, 1999.

Green, R. and D. Newbery, "Competition in the British Electricity Spot Market", *Journal of Political Economy*, v. 100, no. 5 (1992), 929-953.

Green, R., "The Electricity Contract Market in England and Wales", *The Journal of Industrial Economics*, vol.47, March 1999.

Grossman, S. "Nash Equilibrium and the Industrial Organization of Markets with Large Fixed Costs", *Econometrica*, v.49, 1981.

Hart, J. "Imperfect Competition in General Equilibrium: An Overview of Recent Work", in *Frontiers of Economics*, ed. By K. Arrow and S. Honkapohja. Oxford: Basil Blackwell, 1985.

Haurie, A., G. Zaccour and Y. Smeers, "Stochastic Equilibrium Programming for Dynamic Oligopolistic Markets" *JOTA, v.* 66, 1990.

Higle, J.L. and S. Sen, "Duality for Multistage Convex Stochastic Programs" forthcoming in *Annals of Operations Research.*

Joskow, P., E. Kahn, "A Quantitative Analysis of Pricing Behavior in California's Wholesale Electricity Market During Summer 2000", Mimeo, M.I.T., 2000.

Keeney, R.L. and Raiffa H., *Decision with multiple objectives,* New York: John Wiley, 1976.

Klemperer, P.D, and M.A. Meyer, "Supply Function Equilibria in Oligopoly Under Uncertainty" *Econometrica*, v.57, no.6, November 1989.

Knittel, C.R. "Multi Product Firms and Ramsey Pricing: Evidence from US Electricity and Natural Gas Industries", working paper, August 2000.

Kucewicz, W.P., “California’s Dreaming: Energy Policymakers Can’t Defy Economics Forever”, at GeoInvestor.com, 2001.

Kydland, F. and E. Prescott, “Rules Rather Than Discretion : The Inconsistency of Optimal Plans” *Journal of Political Economy*, v. 85, June 1977.

Lave, L., D. Perekhodtsev, “Capacity Withholding Equilibrium in Wholesale Electricity Markets”, Mimeo, Carnegie Mellon University, 2001.

Lockwood, B., "Uniqueness of Markov-perfect equilibrium in infinite-time affine-quadratic differential games" *Journal of Economic Dynamics and Control*, v. 20, 1996.

Marathe, A., C. Barrett, “Market Power in Power Market”, Mimeo, Los Alamos National Laboratory, 2001.

Murphy, F.H., H. Sherali and A. Soyster, “A Mathematical Programming Approach for Determining Oligopolistic Market Equilibrium” *Mathematical Programming,* 1982.

Murphy, F.H., S. Sen and A. Soyster, "Electric Utility Capacity Expansion Planning with Uncertain Load Forecasts" *AIIE Transactions*, v. 14, 1982.

Mussa, M., “External and Internal Adjustment Costs and the Theory of Aggregate and Firm Investment” *Economica,* v. 44, 1977.

NEOS Server, www-neos.mcs.anl.gov

NERA Economic Consulting (www.nera.com) “ A Conversation with Dr.David Salant”, 2002.

Newbery, D., “Competition, Contracts and Entry in the Electricity Spot Market”, *RAND Journal of Economics*, v. 29, no. 4 (1998), 726-49.

Okuguchi, K. and F. Szidarovszky, *The Theory of Oligopoly with Multi-Product Firms,* New York: Springer Verlag, 1999.

Pakes, A. and P. McGuire, "Computing Markov-Perfect Nash Equilibria: Numerical Implications of a Dynamic Differentiated Product Model" *RAND Journal of Economics*, v. 25, Winter 1994.

Perekhodtsev, D., L. Lave, S. Blumsack “The Pivotal Oligopoly Model for Electricity Markets”, Mimeo, Carnegie Mellon University, 2002.

Pitovsky, R. "FTC Chairman Testifies on Electric Power Deregulation: Potential For Consumer Savings is 'Enormous'", at www.ftc.gov, 1997.

Reynolds, S., "Strategic Capital Investment In the American Aluminum Industry" *The Journal of Industrial Economics,* v. 34, 1986.

Reynolds, S., "Capacity Investment, Preemption and Commitment in an Infinite Horizon Model" *International Economic Review*, v. 28, 1987.

Reynolds, S., "Dynamic Oligopoly with Capacity Adjustment Costs" *Journal of Economic Dynamics and Control*, v. 15, 1991.

Rothkopf, M.H. "Control of Market Power in Electricity Auctions", Mimeo, Rutgers University, 2001.

Rudkevich, A., M. Duckworth and R. Rosen, "Modeling Electricity Pricing in a Deregulated Generation Industry: The Potential for Oligopoly Pricing in a Poolco", *Energy Journal*, v. 19, no. 3 (1998), 19-48.

Sen, S., L. Yu and T. Genc, "Decision Aids for Scheduling and Hedging (DASH) in Deregulated Electricity Markets: A Stochastic Programming Approach to Power Portfolio Optimization" *Proceedings of the 2002 Winter Simulation Conference.*

Sen, S., R.D. Doverspike and S. Cosares, "Network Planning with Random Demand" *Telecommunications Systems*, v. 3 (1), 1994.

Sheffrin, A. "Critical Actions Necessary for Effective Market Monitoring." Mimeo, Department of Market Analysis, California ISO, 2001.

Sheffrin, A. "Predicting Market Power Using the Residual Supply Index." Mimeo, Department of Market Analysis, California ISO, 2002.

Taylor, S. "Modeling Financial Time Series," John Wiley & Sons, 1986.

Tracinski, R. W. "California's Antitrust 'Deregulation' of the Power Industry", *Capitalism Magazine*, December 2002.

Vehvilainen, I. "Applying Financial Theory to the Deregulated Electricity Markets". Presentation notes at www.hut.fi/u/iiovo, 2002.

Vehvilainen, I. " Basics of Electricity Derivative Pricing in Competitive Markets", Forthcoming in *Applied Mathematical Finance*, 2002.

WASP-IV: "Wein Automatic System Planning", Version IV, International Atomic Energy Commission, 2000.

Wolfram, C.D. "Measuring Duopoly Power in the British Electricity Spot Market", *The American Economic Review*, vol.89, 1999.