

ALLOCATION AND PRICING OF WATER RESOURCES

A DISSERTATION  
SUBMITTED TO THE DEPARTMENT OF MANAGEMENT SCIENCE  
& ENGINEERING  
AND THE COMMITTEE ON GRADUATE STUDIES  
OF STANFORD UNIVERSITY  
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS  
FOR THE DEGREE OF  
DOCTOR OF PHILOSOPHY

Claire Danielle Tomkins

September 2008

UMI Number: 3332940

### INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

**UMI**®

---

UMI Microform 3332940

Copyright 2008 by ProQuest LLC.

All rights reserved. This microform edition is protected against unauthorized copying under Title 17, United States Code.

ProQuest LLC  
789 E. Eisenhower Parkway  
PO Box 1346  
Ann Arbor, MI 48106-1346

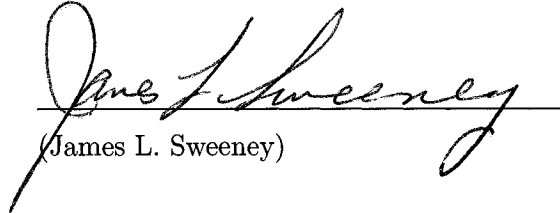
© Copyright by Claire Danielle Tomkins 2008  
All Rights Reserved

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.



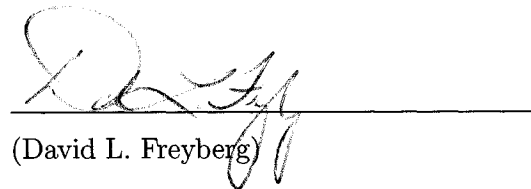
(Thomas A. Weber) Principal Adviser

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.



(James L. Sweeney)

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.



(David L. Freyberg)

Approved for the University Committee on Graduate Studies.



*To my parents*



# Abstract

Public intermediaries responsible for the management of uncertain water supply and demand face a number of decision problems. These include decisions with regard to water procurement and storage, capacity expansion, general operations, investment in demand side management programs, and pricing, among others. This thesis addresses two questions central to the water intermediary's management objective, both related to pricing. The first question is that of an optimal pricing policy in the presence of storage and both demand and supply uncertainty. There is an extensive literature treating the question of optimal public pricing in the presence of joint demand and supply uncertainty with specific application to the electricity sector (peak-load pricing). This literature assumes nonstorability of goods. The presence of storage is a defining feature of the water problem. We treat the pricing problem in the presence of storage, in a joint inventory-price control framework. We characterize the optimal pricing policy and investigate its sensitivity to a number of key model parameters, including the cost of water shortage. The second pricing model developed in this thesis treats option contracting in the water market. Reliance on option contracts is a relatively new phenomenon in the California water market. These contracts can be invoked to avoid costly shortfalls in supply faced by both urban and agricultural water agencies. As such, they can serve an important risk management role in the water market. Option contracts institute temporary transfers with lower transaction cost and lower institutional resistance than permanent transfers. Hence, in addition to providing a risk management tool, they may stimulate the water market by encouraging reliance on short-term water transfers. We price the water option contract from the seller's perspective and from a social planner's perspective. We then investigate the sensitivity of the contract prices and quantity to shifts in a number of contract parameters, including the presence of an outside option for the buyer and possible access to a spot market for water. Finally, we investigate changes to the contract prices and quantity in the presence of multiple buyers and sellers, where the latter requires consideration of strategic interactions between sellers.





# Acknowledgements

My thesis has been conceived and written under the guidance of Thomas Weber. I consider it a great privilege to have had the opportunity to work with Thomas during my time at Stanford. His perspicacity and tireless dedication to the honing of ideas and the pursuit of higher learning are at once inspirational and awing. But beyond that, his integrity, kindness, and sense of humor have made it a delight to work with him.

I am indebted to all of my committee members for the knowledge they have shared and the ideas they have helped instill. Jim, David, and Buzz have all taught me a great deal during my time at Stanford and, in particular, helped foment the ideas in this thesis.

My family. My Grandmother, who resides in the U.K. but corresponds (and calls) to check up on me. My parents, who are endlessly supportive and who I admire so greatly. And my brother, for his humor and his love. Also, Charu and Gomes, who are honorary family members in every sense of that word.

There are a number of good friends I have met during my time at Stanford and who have enriched the PhD process: Mike and Maya, Ayo, Jeffrey, Pedram, Kate, Elisa, Oren, Kenny, Josh, David, Ram, Tzvet, Pete, Lanier, Hal, Clint, Randy and Karen, Ben and Diane, Arjun, Nimmi and Varun, Paul, Beth, and Veena. And of course my very close friend Brandy, who not only kept me company most of my time here but also endeavored to keep me in shape, with her training regimen and various expeditions.

Finally, there are a number of dear friends who have been supportive from afar: Jules (perhaps farthest, in China) and Mel, “the book club” (Anjali, Angel, Karla), friends from high school (Jill, Megan, and Matt), Athena, Tom, Charlotte, and Kim.



# Contents

<b>Abstract</b>	<b>vii</b>
<b>Acknowledgements</b>	<b>ix</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Motivation . . . . .	1
1.2 Research Questions . . . . .	6
1.3 Public Pricing . . . . .	9
1.4 Joint Inventory-Price Control . . . . .	13
1.5 Bilateral Option Contracting . . . . .	16
1.6 Option Contract Competition . . . . .	18
1.7 Application to California Water . . . . .	18
<b>2 Joint Inventory-Price Control</b>	<b>23</b>
2.1 Motivation . . . . .	23
2.2 The Model . . . . .	26
2.3 Optimal Policy and Comparative Statics . . . . .	29
2.3.1 Optimal Policy . . . . .	32
2.3.2 Comparative Statics . . . . .	34
2.4 Application to the Metropolitan Water District . . . . .	40
2.4.1 Background on MWD . . . . .	40
2.4.2 Shortage and Holding Cost Variability . . . . .	44

2.4.3	Shifts in Uncertainty . . . . .	46
2.4.4	Policy Implications . . . . .	47
2.5	Conclusion . . . . .	51
<b>3</b>	<b>Bilateral Option Contracting</b>	<b>65</b>
3.1	Motivation . . . . .	65
3.2	The Model . . . . .	67
3.3	First-Best Allocation . . . . .	70
3.4	Conditions for Contracting . . . . .	75
3.4.1	Contracting without an Outside Option . . . . .	75
3.4.2	Contracting with an Outside Option at a Fixed Price . . . . .	76
3.4.3	Contracting with Joint Market Access: Certain Demand . . . . .	76
3.4.4	Contracting with Joint Market Access: Uncertain Demand . . . . .	79
3.5	Contracting without an Outside Option . . . . .	80
3.5.1	Seller Optimum . . . . .	83
3.5.2	Social Optimum . . . . .	86
3.5.3	Example: Contracting without an Outside Option . . . . .	88
3.6	Contracting with an Outside Option . . . . .	91
3.7	Contracting with Joint Spot Market Access . . . . .	98
3.7.1	Certain Demand . . . . .	98
3.7.2	Uncertain Demand . . . . .	100
3.7.3	Example: Contracting with Joint Spot Market Access . . . . .	103
3.8	Application to the California Water Market . . . . .	108
3.8.1	Background . . . . .	108
3.8.2	Contract Parameters . . . . .	110
3.8.3	Welfare Gains Under Option Contracting . . . . .	113
3.8.4	Price Trends . . . . .	114
3.8.5	Future Contracting . . . . .	117

3.9	Summary	117
3.10	Appendix	120
<b>4</b>	<b>Option Contract Competition</b>	<b>129</b>
4.1	Motivation	129
4.1.1	Related Literature	130
4.1.2	Outline	134
4.2	Multiple Buyers	134
4.3	Multiple Sellers	136
4.3.1	Two Sellers with Complete Information	136
4.3.2	Two Sellers with Incomplete Information	139
4.3.3	Multiple Sellers with Incomplete Information	145
4.3.4	Distribution of Gains from Trade	147
4.4	Summary	150
4.5	Appendix	152
<b>5</b>	<b>Conclusion</b>	<b>157</b>
5.1	Contributions	157
5.2	Application of the Models	160
5.2.1	Price-Inventory Control at MWD	160
5.2.2	Option Contracting in the California Water Market	163
5.3	Areas for Future Work	167
5.3.1	Inventory-Price Control	167
5.3.2	Contracting	169



# List of Tables

2.1	Demand Estimation Parameters . . . . .	43
2.2	Parameter Values for MWD . . . . .	43
2.3	MWD's Incremental Interruption and Conservation Plan, Stages I-IV . . . . .	62
2.4	Unbundled Water Rates for Financial Year 2006 . . . . .	63
2.5	Historical Water Rates: 1986–2006 . . . . .	64
2.6	Shortage Cost Estimates by Sector . . . . .	64
3.1	Rice Production Data . . . . .	111
3.2	MWD's Unbundled Wheeling Charges (per af). . . . .	112
3.3	Surplus Division and Social Welfare Gain with Option Contracts . . . . .	114
3.4	Historical Contract Types and Transfer Volumes. . . . .	116
3.5	Sensitivities to the Price of Rice (in millions of dollars). . . . .	118





# List of Figures

1.1	Water Intermediary's Operational Environment . . . . .	4
1.2	Joint Price-Inventory Control . . . . .	15
1.3	Bilateral Option Contracting . . . . .	17
2.1	A Price Increase Under Supermodularity . . . . .	31
2.2	20-Year Demand and Supply . . . . .	41
2.3	Supply-Demand Correlation . . . . .	42
2.4	Sensitivity of Policy to Holding Costs ( $r = \$2,500$ ) . . . . .	45
2.5	Sensitivity of Policy to Holding Costs ( $r = \$0$ ) . . . . .	45
2.6	Sensitivity of Policy to Shortage Cost . . . . .	46
2.7	Sensitivity of Policy to Shifts in the Mean of Demand . . . . .	48
2.8	Sensitivity of Policy to Pure Demand Uncertainty . . . . .	48
2.9	Sensitivity of Policy to Pure Supply Uncertainty . . . . .	49
2.10	Sensitivity of Policy to Joint Uncertainty . . . . .	49
2.11	Sensitivity of Payoff to Uncertainty . . . . .	50
2.12	Sensitivity of Policy to the Correlation Coefficient $\rho$ . . . . .	51
2.13	Sensitivity of Payoff to the Correlation Coefficient $\rho$ . . . . .	51
2.14	Sensitivity of Policy to Profit vs. Welfare Maximization . . . . .	52
2.15	Sensitivity of Policy to the Correlation Coefficient $\rho$ (Welfare Maximization) . . . . .	52
3.1	Expected Surplus . . . . .	74
3.2	Comparison to the First-Best Allocation . . . . .	87
3.3	Sensitivity of the Contract Level and Prices to Expected Market Price . . . . .	89

3.4	Sensitivity of Buyer's and Seller's Payoffs to Expected Market Price . . . . .	89
3.5	Comparison of Contract-Level and Payoff Under Profit Maximization and Social Welfare Maximization . . . . .	90
3.6	Deadweight Losses Under the Seller-Set Optimum . . . . .	90
3.7	Profit Analysis . . . . .	95
3.8	Sensitivity of Option Price to the Price of the Outside Supply Option . . . . .	96
3.9	Optimal Contract Level . . . . .	96
3.10	Expected Purchases from the Outside Option . . . . .	97
3.11	Optimal Prices with Joint Spot Market Access . . . . .	104
3.12	Contract Level and Profit Sensitivity to the Market Friction Coefficient $\tau$ . .	104
3.13	Payoff and Social Welfare Sensitivity to the Market Friction Coefficient $\tau$ (Under Profit Maximization) . . . . .	105
3.14	Order Level and Social Welfare Sensitivity to the Market Friction Coefficient $\tau$ (Under Social Welfare Maximization) . . . . .	105
3.15	Social Welfare Comparison Under Profit Maximization vs. Social Welfare Maximization . . . . .	106
3.16	Social Welfare Optimal Prices . . . . .	106
3.17	Payoffs Under Social Welfare Maximization . . . . .	107
4.1	Aggregate Demand . . . . .	135
4.2	Equilibrium Regions: Complete Information . . . . .	138
4.3	Equilibrium Price Regions ( $\pi_H^*$ with $c = 0$ ) . . . . .	142
4.4	Equilibrium Price Regions ( $\pi_L^*$ with $c = 0$ ) . . . . .	143
4.5	Price Comparison: $p_H^*$ and $p_L^*$ . . . . .	144
4.6	Optimal Bidding Function (for $N = 2$ ) . . . . .	148
4.7	Optimal Bidding Function (for $N \in \{2, \dots, 7\}$ ) . . . . .	149
4.8	Cdf of Surpluses (for $N = 2$ ) . . . . .	149
4.9	Cdf of Surpluses (for $N = 3$ ) . . . . .	150
4.10	Capacity Regions for $r_i > r_j$ . . . . .	153

5.1	Current Rate-Setting Practice . . . . .	160
5.2	Dynamic Rate-Setting Process . . . . .	161
5.3	Steps in the “Cost of Service” Rate-Setting Methodology . . . . .	161
5.4	Categories for Cost Allocation . . . . .	162



# Chapter 1

## Introduction

### 1.1 Motivation

The provision of freshwater to consumers generally falls in the province of the public, rather than the private, sector. There are several possible explanations as to why this is the case. Bös (1985) argues, for instance, that government tends to oversee the provision of services that are essential to public well-being. When a service is essential, the risk of a catastrophic market failure justifies, or motivates, government intervention. His examples include government oversight of the gas, electricity, and water industries. Another explanation is that the water sector is a natural monopoly, where the large infrastructure costs and the decreasing returns to scale nature of the business imply the optimality of a single operator. The government then regulates the monopoly to protect consumers from monopolistic pricing. Yet another explanation is limited interest on the part of the private sector in the task of water provision, leaving the government no choice but to undertake the responsibility. These reasons may have more or less explanatory power depending on the setting under examination. For instance, the management of water services by the public sector in socialist countries is consistent with the prevailing philosophical doctrine regarding the government's responsibility to provide basic services to all citizens. In capitalist countries, the explanation may owe more to the existence of natural monopoly than a conviction that the government *should* provide all basic services. In some parts of the developing world there may be insufficient capital and coordination in the private sector to undertake the task of water supply, and government oversight is therefore the default option. While the explanations as to why may vary, the provision of freshwater by government entities is the rule not the exception

worldwide.

The public enterprise (or regulated private enterprise) charged with the provision of water to consumers faces a supply-demand management task. Abstracting from reality for a moment, we consider the enterprise's task in a world where supply is fixed, or known.<sup>1</sup> When there is a predetermined amount of water available at every point in time, the supply-demand management task is reduced to one of allocation. Given the supply currently available, *how should the water be distributed amongst consumers?* If individuals' willingness to pay (WTP) is known, then the government can compute the optimal quantity to deliver to each household. Alternatively, the government can compute the market-clearing price given aggregate supply and demand information. Under perfect information, the price and quantity instruments at the government's disposal perform equivalently from a welfare standpoint. That is, the government can either calculate the optimal quantity to deliver to each household and then levee a fixed fee to recover delivery costs *or* it can calculate the market-clearing price based on the supply cost and the aggregation of individual WTP, and in both cases it can achieve the welfare optimum. We note that when quantities are assigned and costs recovered via a fixed fee, the fixed fee must be nondiscriminatory to achieve welfare optimality. This can be achieved by charging each customer, or class of customers, a separate fee in accordance with WTP.

Under uncertainty, price and quantity instruments do not perform equivalently. For instance, we consider the following setting: assume that supply is no longer fixed but, rather, is provided by a private contractor. The private contractor will build a reservoir, and the total volume of water available for consumption will be equal to the capacity of the reservoir. The public enterprise faces the following decision: it can either announce a price for water, which will form the basis for the private contractor's reservoir capacity decision, *or* it can mandate a given reservoir capacity, allowing the price to be determined by market-clearing conditions given the realization of demand. The cost of building the reservoir is uncertain. The exact benefits to the consumers of water supply are also uncertain. If the objective is surplus maximization, which instrument does the enterprise select? In a seminal paper, Weitzman (1974) provides conditions under which reliance on one instrument over the other (price vs. quantity) is preferable from a social welfare standpoint.

The context in which Weitzman poses the question is that of an entire economy, in which various firms and consumers respond to the price or quantity announcement by the government (referred to as "the centre"). Nonetheless, if we view the public enterprise's decision

---

<sup>1</sup>We also temporarily abstract away from the possibility of demand-side management programs such as subsidized conservation.

as a choice regarding which instrument to use, the Weitzman analysis is instructive. When either of the following two conditions hold, quantities are the preferred instrument: (1) the benefit function is sharply curved or (2) the cost function is close to linear. As Weitzman emphasizes, the result hinges on the *relative*, not absolute, curvature of the cost and benefit functions. The implication of (2) is that a small adjustment in the announced price will lead to gross over or under-production. In case (1), when the benefit function is sharply curved and, hence, the marginal benefit of an additional unit of the good is high, the loss from running low on supply will also be high. In this case, it is preferable to set quantities and guarantee that supply does not run short – a conservative policy. In our contrived example, if the water intermediary suspects that the benefit of water supply is steeply curved at the optimum, it should fix capacity and allow the price to be determined by market-clearing conditions once demand realizes.<sup>2</sup>

As acknowledged by Weitzman, the question of price vs. quantity necessarily transports us to the realm of second best. There is no reason *a priori* to eliminate consideration of other instruments, including state-contingent price schedules. Rather, the interest in the problem arises from the observation that the two simple instruments are often employed in practice. By increasing the number of controls – for example, by considering the *joint control* of capacity and price above – we can improve on the Weitzman outcome.

The supply-demand management problem faced by a water intermediary encompasses a number of decisions, or controls, related not only to capacity and water tariffs but also storage operations, conservation programs, and supply contracting, to name a few. Figure 1 depicts the water intermediary's management milieu. The natural supply of water is that which comes directly from the environment, e.g., from rainfall and runoff. The natural water supply is an exogenous variable. That is, we assume that the government (however far-reaching its powers may be in other areas) cannot alter the hydrologic cycle.<sup>3</sup> Water from the hydrologic system can enter the managed system directly as, say, inflow to a canal or aqueduct, or through storage, which includes recharge of aquifers. An intermediary's supply can be augmented through supply contracts with other (private or public) intermediaries. These supply contracts may be long-term supply contracts or one-shot deals, i.e., option contracts. For instance, there may be a number of regional water districts coordinating supplies to different users. The presence of multiple public enterprises, each managing separate

---

<sup>2</sup>Also of relevance is Weitzman's observation that in the case where the the marginal cost is steeply increasing at the optimum, as in the case of fixed costs, the choice of prices vs. quantities results in approximately the same output.

<sup>3</sup>Interestingly, in light of climate change, this assumption as a characterization of the long-term relationship between government action, consumer habit, and the hydrologic cycle is not strictly accurate.



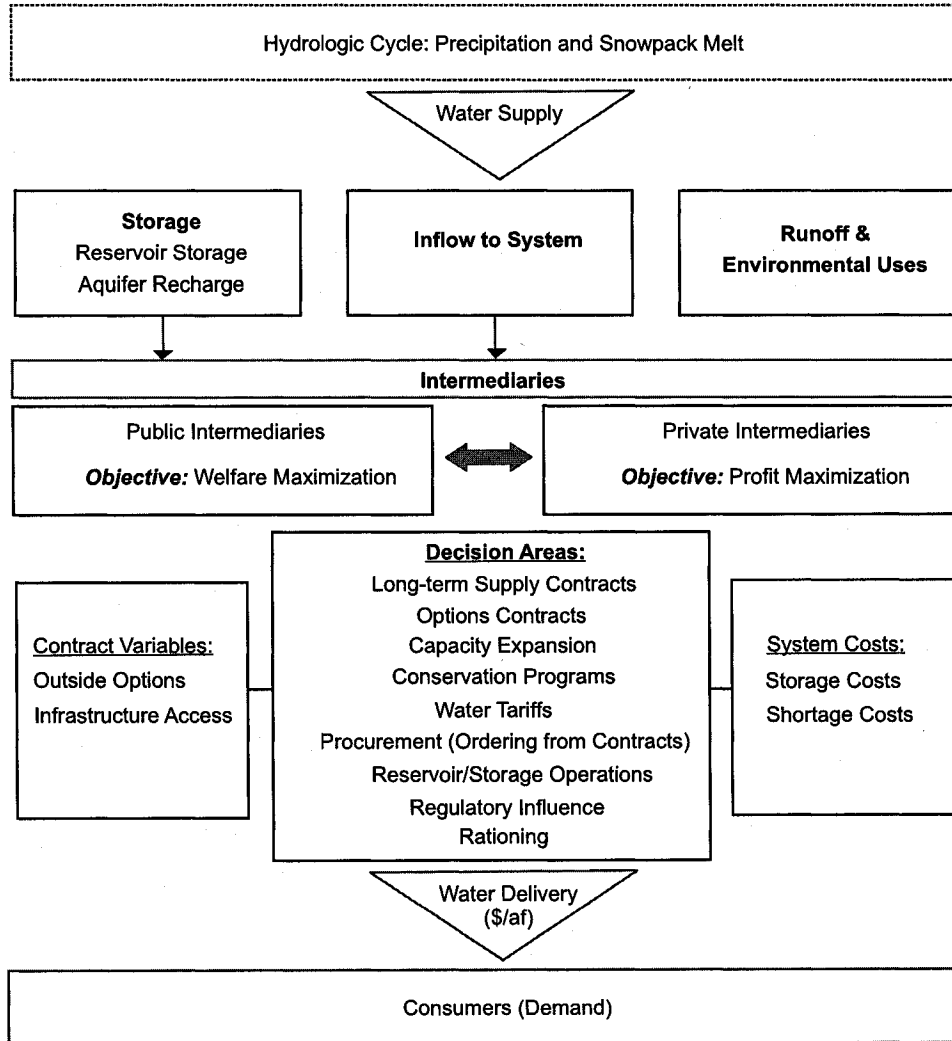


Figure 1.1: Water Intermediary's Operational Environment

water supplies, creates the potential for the negotiation of interregional, i.e., interdistrict, interstate, or even international, water transfers.

The decisions the intermediary faces can be loosely grouped into supply-side management decisions and demand-side management decisions. On the supply-side, the intermediary faces decisions regarding how much storage to build (a capacity decision), how much water to store and/or extract in each period (an operational decision) given capacity, as well as contracting and ordering, or procurement, decisions. Contracting for supply may involve negotiating water rights purchases or agreements for future water delivery – there are currently no spot markets for water in the world. The negotiated transfer of a permanent water right, also referred to as a statutory transfer, has been the focus of much of the literature on and activity in developing water markets to date.

On the demand-management side, the most prominent decision is that regarding water tariffs – both the structure and level of tariffs. The intermediary can also encourage and/or subsidize conservation measures. This has become common practice by both electricity and water utilities. In periods of shortage, the intermediary can impose a rationing policy.

A wholistic approach to the determination of the intermediary's optimal supply-demand management policy would take into consideration all of the possible decisions outlined in Figure 1. It is worth pausing here to note, however, that not all of these decision variables can be adjusted in the short-run. For instance, capacity is generally fixed in the short-run. Supply contracts may or may not be renegotiated depending on the length of the time horizon. That is, certain variables can be dynamically adjusted – others cannot. This leads to a natural division in the types of optimization problems that we might consider for our water intermediary, depending on the relevant time horizon.

The research questions that we outline in the next section and proceed to examine in the remaining chapters of this thesis are addressed under a number of simplifying assumptions – most notably the assumption that the intermediary's other decision variables are fixed. For example, when considering optimal rate-setting and ordering policies, we will assume that storage capacity, supply contracts, and the distribution of demand outcomes are all fixed. This assumption facilitates the construction of analytically tractable models, but its justification does not rest on that fact.

The argument in favor of our simplifying assumptions is a direct appeal to that offered by Weitzman at the start of his analysis. Namely, that these problems are of interest precisely because they are common in practice. In practice, delegation within an organization, timing of decisions, and imperfect foresight, not to mention limitations on time and computational

power, mean that decisions are often made independently. Furthermore, as just mentioned, a number of variables remain fixed in the short-run. Our models address optimal decision-making given these practical limitations.

The rest of this chapter is organized as follows. In Section 1.2 we introduce and motivate our three main research questions. Then in Section 1.3 we review the relevant public pricing literature. In Section 1.4, we introduce the dynamic inventory-price control model and discuss the relevant OM literature. Sections 1.5 and 1.6 discuss contracting for water in a bilateral and multilateral environment, respectively, and again review the relevant literature. Finally, in Section 1.7 we provide a detailed background on the example we will use throughout this text to illustrate the applicability of our models, that of the existing California water system.

## 1.2 Research Questions

This thesis addresses three questions regarding optimal supply-demand management by a water intermediary facing supply and demand uncertainty. The underlying motivation of the analysis is the possibility of costly water shortages. Hence, in Chapter 2, when we ask, *what is the efficient pricing and supply procurement policy by a large water intermediary?*, the answer depends on the shortage cost of water. In Chapter 3 and Chapter 4, where we examine option contracts for water as a way to avoid costly shortfalls, asking, *what is the optimal structure of these contracts?*, the answer again depends critically on the value of these avoided water shortages.

These questions are arguably ones of great importance in light of the rising value (or cost) of water worldwide, coupled with the increasing incidence of water shortage. A number of pressures exerted on existing water systems can lead to water shortages: prolonged drought, changes in the timing of rainfall/runoff patterns (due to global climate change), increasing demand (due to population growth or economic expansion), and environmental constraints, to name a few. In developing countries, these conditions are often exacerbated by limited infrastructure.

Our first question is with regards to the intermediary's pricing policy. If supply is fixed, as it presumably is in the short-run (where the short-run could be defined such that this assumption holds, e.g., this instance), then the decision sphere is limited to ordering and demand-side management policies. To be clear, by assuming short-run supply is fixed, we are *not* assuming that the amount of supply is deterministic. Rather, the distribution of possible supply outcomes as dictated by the environment *and* the underlying contractual

agreements is fixed. Our first research question considers the optimal dynamic joint ordering and rate-setting policy for the intermediary

**Research Question 1** *Given fixed short-run supply contracts and a given storage level, what is the optimal joint rate-setting and supply procurement policy by a water intermediary facing both supply and demand uncertainty?*

We abstract from the possibility of various alternative demand-side management programs and focus on the optimal per-period price. It is common practice for a public water intermediary to charge a uniform tariff to its consumers. The integration of the price and ordering decision is not common practice, owing in large-part to the fact that rates are often set far in advance and supply procurement is an ongoing process. Interest in efficiency gains from real-time, or dynamic, pricing, which would allow coordination of these two decisions, motivates our examination of this problem. The possibility of additional reliance on supply contracts by water intermediaries in the future provides a secondary motivation. As discussed in more detail in the following section, the question of optimal pricing and ordering *in the presence of storage* has not been formally addressed in the literature to date. In Chapter 2, we develop a model of the intermediary's joint inventory-price control problem and derive an optimal policy under two-sided uncertainty.

In Chapter 3, we introduce a bilateral contracting model for water options. The motivation for our study of option contracting is two-fold. First is a long-standing interest in water markets as a possible avenue for achieving efficient reallocation of water resources. After a lengthy (several-decade-long) focus on the transfer of permanent water rights, also referred to as statutory transfers, we observe increasing reliance on temporary transfers in developing water markets. The possibility of interregional water transfers exists in most parts of the world and has spurred wide interest in water markets to reallocate water and avoid costly shortfalls in supply. There are as yet no spot markets for water. There are, however, an increasing number of examples of contractual water transfers. While there are increasing number of documented contracts in the market, there are not, to the best of our knowledge, existing contracting models to provide insight into optimal contract structure and associated efficiency issues. It is our aim, in Chapter 3 and Chapter 4, to address these issues.

The second motivation comes directly from the electricity sector, where reliance on option contracts to mitigate supply shortages has been widely and successfully implemented. Typical contracts in electricity markets can be grouped into two categories, options or forwards.

Many variants of these basic contracts exist, e.g., extendable options. Both types of contracts are of interest in the context of water.

**Research Question 2** *What is the structure of the optimal bilateral option contract given downstream demand uncertainty, uncertain valuations on outside options for the buyer and the seller, and the possibility of infrastructure failure?*

We address question two in several alternative contracting environments. First we consider the case where the buyer does not have an outside supply option and is, therefore, beholden to the seller. We then relax this assumption and consider contracting in an environment where there is an outside supply option available at a fixed price. Finally, we consider contracting in the presence of joint access to a spot market for water, a possible future reality.

In Chapter 4, we extend our analysis to consider the entrance of additional buyers and sellers into the contracting market, asking,

**Research Question 3** *How does the presence of additional buyers and sellers in the contracting market impact the structure of optimal option contracts?*

With regards to question three, we specifically consider changes in the contract structure for the following two cases: a single seller and multiple buyers, and a single buyer and multiple sellers. Strategic interactions in the case of multiple sellers significantly change the structure and analysis of the option contracting model.

Finally, to address these questions satisfactorily, we must identify our objective. A natural tendency in public economics is to maximize allocative efficiency, i.e., to maximize social welfare, where social welfare is defined as the sum of consumer and producer surplus. The ideal represented by this objective can hardly be faulted. (Once surplus is maximized, distributional objectives can be met through transfers.) The reality, however, is questionable. In a review of public pricing, Bös (1985) posits several alternative objectives used in practice by regulated public enterprises, including pricing to achieve specific distributional objectives of government, pricing to achieve politicians' aims, e.g., to be re-elected, and pricing to achieve managers' aims, e.g., to maximize budgets and thereby gain influence.

In the models we develop herein, we consider outcomes under both profit-maximizing and social-welfare maximizing behavior on the part of the enterprise in question. In the context of the pricing model for a water intermediary in Chapter 2, we posit that the profit-maximizing

objective is representative of bureaucratic aims to maximize influence. In the contracting models developed in Chapter 3 and Chapter 4, profit-maximization on the part of the seller seems a natural assumption, whereas the welfare-maximizing solution is of clear interest from a normative perspective. In general, we will compare outcomes under the two differing objectives.

## 1.3 Public Pricing

Our first research question, that of optimal rate-setting, has been discussed in detail in the public pricing literature, warranting a short review here. In particular, the literature on 'peak-load' pricing treats optimal pricing under demand uncertainty. As we discuss below, this literature incorporates two major assumptions that make it inappropriate for the water allocation problem under consideration. This motivates our introduction of joint inventory-price control for water.

Beginning with Ramsey (1927), the public economics literature has developed the topic of public pricing in detail. The prevalent models assign price-setting power to the governing board of the public enterprise in charge of allocation. As noted above, with regard to the board's objective, the tradition is welfare maximization, although a number of other objectives have been posited and explored. The motivation for public pricing models comes from the observed number of goods and services for which the government assumes responsibility. Bös (1985) notes that the provision of *essential goods and services* – goods and services without which the underlying infrastructure of the economy is threatened – is often overseen directly or indirectly by the government (presumably because market failure in these sectors could have catastrophic consequences for society). The list of essential goods and services includes electricity, gas, water, airlines, railroads, tollways, telephone, and postal service, among others. (There are of course examples of publicly priced non-essentials, and Bös provides one with respect to cigarette-pricing by French and Austrian monopolies.) Where confidence in the market is strong, some of these sectors have undergone privatization on the grounds that the private sector can ultimately achieve greater efficiency through competitive pricing. For example, electricity markets in California are operated in the private sector, with some government oversight.

The Ramsey model solves a welfare maximization problem subject to three constraints on the behavior of the economy: (1) consumption equals production, i.e., market-clearing conditions, (2) the public sector produces efficiently, and (3) a revenue-cost constraint must be

satisfied. Ramsey pricing can ultimately dictate prices above or below marginal cost. In the simplified case where cross-price elasticities between goods are ignored, Ramsey pricing can be characterized by the inverse elasticity rule, where the price-cost margin of a good is proportional to its own-price elasticity of demand. One important implication of this rule, as observed by Bös, is that prices optimally lie above marginal costs when a public enterprise faces a break-even constraint and operates using a production technology exhibiting increasing returns to scale. Conversely, if the enterprise is constrained to offer prices below a certain level, i.e., the revenue-cost constraint is low (even negative), then optimal prices may fall below marginal costs. At no point are prices exactly equal to marginal cost in this model.

Yet, marginal cost pricing has been advocated in many public settings and adopted in some cases. What, we might ask, is the theoretical justification for this? The optimality of the marginal-cost pricing rule is a direct result of the relaxation of the revenue-cost constraint. This points to the potential problem encountered under marginal-cost pricing: deficits. The solution to which is the introduction of two-part tariffs, or non-linear pricing, whereby the deficits can be recouped via a fixed fee. These two-part tariffs are generally welfare optimal. Although, as Oi (1971) and Ng and Weisser (1974) discuss, their optimality is compromised when consumer participation is considered endogenously. For example, a fixed-fee that excludes lower-income users who would have been willing to pay the marginal cost rate is non-optimal. This can in theory be addressed by differentiated customer rates.

The Ramsey model received renewed interest with publication of Boiteux's 1956 article addressing rate setting in a budget-constrained public utility. The Boiteux model extends the Ramsey model and has formed the basis of much discussion surrounding public pricing. Boiteux's particular interest at the time was the management of the French national electricity sector, of which he was head. The model he proposed provided the justification for rate reform in this sector.

These early models did not address the question of allocation under uncertainty. Demand uncertainty was later incorporated into the Boiteux framework and has been treated in a succession of models that comprise the peak-load pricing literature. The problem of peak-load pricing is motivated by the electricity sector. Demand for electricity typically experiences several "peaks" diurnally and also seasonally. (For instance, Lord John Browne, the former director of British Petroleum (BP), humorously observed in a lecture at Stanford that electricity consumption in the U.K. consistently experiences a peak at around four in the afternoon when the citizenry boil water for their afternoon cup of tea!) In general, demand fluctuates cyclically.

The time-dependence of demand is captured by the introduction of separate periods in the model and a distribution on demand realizations in each period. Demand in each period is notably assumed to be independent, an assumption that we also adopt in our modelling approach. The peak-load pricing literature has evolved to treat uncertain supply and to integrate a second decision into the optimization framework, that of generation or capacity-expansion. Crew et al. (1995) provide a review of the literature.

The peak-load pricing literature has expanded in several directions since the comprehensive survey provided in Crew et al. (1995). Hirschberg (2000) observes that time of day (TOD) pricing, a form of peak-load pricing, is not efficient unless the patterns of TOD *substitution* are known. This is the shifting-peak problem. Hirschberg presents a statistical method for estimating substitution patterns. Fischer and Serra (2002) introduce plant indivisibilities, improving the model's realism by accounting for the fact that typical plants have a minimum operating level. Arellano and Serra (2007a) examine the exercise of market power when firms face price regulation, subject to peak-load pricing rules, but retain decision-making power regarding their generation portfolio. Arellano and Serra (2007b) extend the model to include transmission costs, and they then investigate how these transmission costs should be optimally allocated, e.g., to generating companies or to consumers. Finally, Pineau and Zaccour (2007) relax the assumption of independent peak and base-load segment demand in the model and examine the dependence of optimal decision-making on the cross-price elasticities.

Several other studies have considered the implementation of peak-load pricing. Doorman (2005) explores capacity subscription on the part of consumers as an approach to improving system reliability. Two studies have investigated household response to peak-load pricing: Herter (2007), in the U.S., and Matsukawa (2001). Yano and Newman (2007) provide an application of peak-load pricing to the express package delivery sector.

On the pricing side, Skiera and Spann (1999) present evidence that firms can use prices effectively to compensate for suboptimal capacity in a peak-load pricing framework. Several other recent studies have contributed to the literature on public pricing and, in particular, the treatment of welfare optimal linear pricing by a monopolist and general nonlinear pricing. De Borger (2000) investigates optimal nonlinear pricing, or two-part tariffs, in a discrete choice model, noting that most prior investigations have assumed a, "continuum of consumers characterised by a one-dimensional taste parameter." Garcia et al. (2001) develop a duopoly model of Bertrand price competition between hydro generators in the electricity sector and the overall impact of such competition on system reliability, under regulatory price caps. Cowan (2003) considers how to optimally adjust a regulated monopoly's price



under demand and cost uncertainty. He considers optimal price adjustments under a number of commonly adopted pricing schemes, including average cost pricing, price caps, and different pass-through rates for marginal costs.

With respect to our water allocation problem, the peak-load pricing literature suffers from one glaringly inaccurate assumption, that of non-storability.<sup>4</sup> For an example of an explicit treatment of storage in a dynamic pricing model, also incorporating demand uncertainty, we turn to a literature that takes its motivation from an entirely different context, the retail sector. As is discussed in Section 1.4 below, the operations and management (OM) literature provides examples of dynamic price and order, or procurement, control models under uncertainty that provide a framework for addressing our water allocation question. The extension of existing models to incorporate supply uncertainty is part of our exposition in Chapter 2.

As noted earlier, there are instances of publicly priced goods shifting into the private sector – the privatization of electricity markets in California provides one such example. Under privatization, generation and distribution enterprises are owned by separate (private) entities. The coordination of supply and demand by an intermediary in the electricity sector then entails contractual relationships with generators (suppliers), as well as dependence on a spot market. This situation is analogous to that faced by a water intermediary when there exist other public or private enterprises from which the intermediary could buy or sell water, e.g., when there is the potential for water transfers.

The privatization of the electricity sector resulted in the creation of a market for electricity. However, privatization is not a prerequisite for such a market. In a situation with many independent *publicly-owned* enterprises, the desire to trade could also generate a market. This is the scenario that currently exists in the state of California (and, indeed, much of the western U.S.): there are literally hundreds of public enterprises in the state managing local water supplies, each with differing water entitlements and differing values (or costs) of water supply, suggesting the possibility of mutually beneficial trade. There is no spot market for water, but districts do negotiate water transfers or, in some cases, purchases of permanent water rights.

The possibility of a water market suggests an expanded view of the water allocation problem. The intermediary introduced at the beginning of this section is not operating in isolation,

---

<sup>4</sup>The assumption of non-storability does not hold in general even in the electricity sector if one is also considering hydropower. The need for pricing models that treat electricity storage is likely to be the rule more than the exception as alternative energy sources such as wind, with subsurface storage potential, and plug-in cars and PVs, both with battery storage potential, enter the generation mix in larger quantities.

nor is his supply purely a function of the hydrologic conditions and his own infrastructure investments, i.e., storage and generation capacity. The intermediary's supply-side is actually comprised of many potential suppliers in the form of other water districts willing to trade. The pricing model we develop for a water intermediary in Chapter 2 embeds the decision regarding the volume of water available under contract. That is, we allow the intermediary to optimally "order" from his existing supply contracts (under the assumption that he has flexible supply contracts that do not require him to take water when he has no use for it.) The bilateral contracting model developed in Chapter 3 treats the contracting decision explicitly, deriving optimal prices and contract quantities for bilateral contracting between water districts. Extensions to a multilateral setting are discussed in Chapter 4.

## 1.4 Joint Inventory-Price Control

The possibility of storage differentiates the water rate problem from the more widely studied problem of rate-setting under uncertainty in the electricity sector. As noted, the models developed to investigate optimal peak-load pricing in the electricity sector rely on the assumption that the good being priced is a nonstorable commodity. In a different context entirely – that of the retail industry – the operations and management (OM) literature has evolved to treat price control in the presence of storage and demand uncertainty.

Specifically, there are a number of joint price-inventory control models with relevance to the water sector. These models originate in the inventory control literature, beginning with Arrow et al. (1951). Inventory control generally refers to the establishment of optimal ordering policies by an intermediary, given current inventory levels. In a model with multiple periods, this amounts to addressing the question *how much should I order given my current inventory level?* in each period. For instance, a shoe retailer decides how many pairs of a given brand of shoe to order this month, given the number of pairs he still has leftover from last month. The factors influencing the retailer's decision include the per-unit charge for his order, the cost of storing excess pairs of shoes (if any), and his projected demand (including the cost to him of potentially not meeting this demand, if he under-orders).

The inventory control literature can generally be subdivided based on its treatment of static versus dynamic models and models with and without uncertainty, where variants of all four combinations exist. It is worth noting that uncertainty refers to pure demand uncertainty. We are only aware of one study that incorporates uncertain supply, or yield (Li and Zheng, 2006). This is perhaps somewhat surprising given that there is presumably some positive

probability of non-delivery of orders. It is also worth noting that an additional dimension of the literature is the treatment of price control, in addition to inventory control. This joint treatment originates early in the literature, with Whitin (1955). Finally, we note that there are number of papers that again dissociate the two and treat dynamic price control, e.g., Gallego and van Ryzin (1994).

Arrow et al. (1951) introduce both a static and a dynamic model of inventory-control model, with an exposition that incorporates uncertainty. The static version of the inventory-control model under certainty yields an optimal order level, or lot size, and, hence, is often referred to as the “lot-sizing model” in the literature. With the addition of demand uncertainty, the static model is the familiar newsvendor model, the solution to which is an optimal newspaper order level given costly orders and uncertain daily demand. Whitin (1955) introduces the price variable to the static models, treating both the case of certain and uncertain demand. Mills (1959) contrasts a monopolist’s joint production and pricing decision under uncertain demand with that under certain demand, finding that for the case of constant marginal cost, optimal prices under uncertainty will be lower. In the general case, one can find cases when it is optimal to increase prices in comparison to the riskless case. Zabel (1972) is the first to treat dynamic price-inventory control, under uncertainty and treats two specific cases, that of multiplicative and additive demand uncertainty. Both Thomas (1974) and Thowsen (1975) develop dynamic finite-horizon models of joint order/production and price decisions. These early contributors laid the framework for what has become an expansive and nuanced literature, treating a host of variations, including generalized demand specifications (versus additive or multiplicative demand uncertainty), backlogged prices equal to fill-period prices, and setup costs, to name a few. Petruzzi and Dada (1999) offer a review of the newsvendor problem with the integrated pricing decision. Talluri and van Ryzin (2004) offer a comprehensive review of revenue management.

The area continues to expand. Petruzzi and Dada (2002) incorporate learning into the model, assuming that a parameter of the demand function is unknown to the intermediary initially. The intermediary monitors demand response, thereby learning, and updates its price-inventory decisions accordingly. Lin (2006) also treats learning in a dynamic pricing model with an imperfect demand forecast. Gallego and Ozer (2001) consider the introduction of advance demand information, a model in which demands are revealed before the fill, or delivery, date and generally with different delivery date. Rajaram and Yin (2007) consider a finite horizon model where demand is no longer assumed to be independent across periods and is, instead, modelled as a Markov chain.

In the context of water, the intermediary’s inventory-price control problem is one regarding

how much water to order, or buy, from his existing supply contracts, given the amount of water he currently has in storage. These supply contracts can be thought of quite generally in terms of a maximum amount of water that the intermediary can obtain in a given period (from a given source). A supply contract as we intend it need not be a contract negotiated between the intermediary and another party. It could, for example, be a water right that entitles the intermediary to a certain amount of water to be “delivered” to him in the form of snowmelt runoff available for diversion from a major river system. In fact, an intermediary is likely to hold supply contracts both in the form of water rights *and* contracts negotiated with other parties, i.e., other water districts. There are four sources of supply (apart from storage) available to a water intermediary, each of which can be viewed in general terms as a supply contract with an associated delivery cost: (1) runoff (from precipitation or snowmelt) for which the intermediary holds a water right (2) groundwater (for pumping), again for which the intermediary holds a right, (3) generation (using desalination or reuse technology), and (4) supply contracts with other water districts.

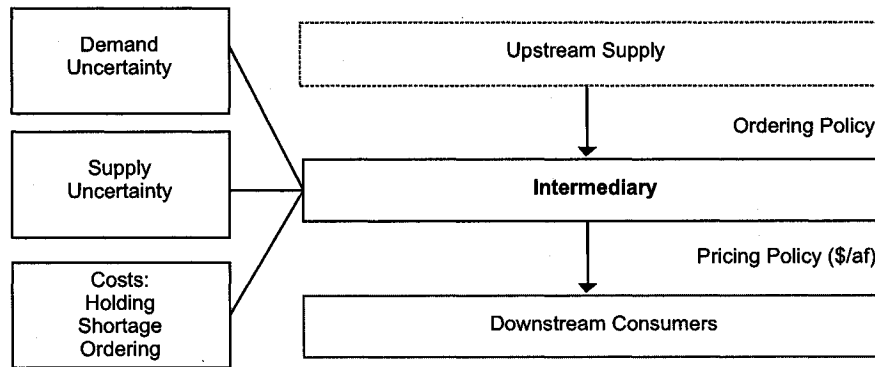


Figure 1.2: Joint Price-Inventory Control

Under the assumption that supply contracts are flexible, i.e., that the intermediary decides how much supply to fill from each contract and is not forced to take water in periods in which he doesn't need it, the analogy to the inventory-control problem is direct. In each period, the water intermediary decides how much water he would like to procure based on his existing supply contracts and the factors mentioned above: the per-unit cost of the supply, the cost of storage, and the cost of shortage, i.e., the cost of not meeting downstream demand.

The model developed in Chapter 2 extends the joint inventory-price control model introduced by Federgruen and Heching (1999), henceforth referred to as F&H. The F&H model is a finite horizon single-product periodic-review model with ordering costs proportional to the quantity

ordered and no setup costs. F&H establish that the optimal policy under demand uncertainty is a base-stock-list-price policy, commonly referred to as an  $(s, S, P)$ -policy. The model developed in Chapter 2 extends the F&H model to treat two-sided (demand *and* supply) uncertainty. We are only aware of one other model that incorporates supply uncertainty. A recent paper, Li and Zheng (2006) discussed yield, or supply, uncertainty.

In addition, we dispense with the standard “backlogging assumption,” whereby all unmet demand is carried over to, and met, the next period. (Unmet demand for water cannot generally be filled in a subsequent period, as might be the case for retail goods.) The no-backlogging model, or “lost sales” model has been examined by Chen et al. (2006) and, prior to that, by Polatoglu and Sahin (2000). Chen et al. (2006) present a “lost sales” model with fixed costs, also referred to as setup costs. They show that with the introduction of a number of additional assumptions, and using arguments based in part on Petruzzi and Dada (2002), it is still possible to establish the optimality of a base-stock-list-price-policy in the lost sales model. As we demonstrate, however, in the relaxed model (without these additional assumptions), their results do not hold generally.

## 1.5 Bilateral Option Contracting

The existence of a water market changes the intermediary’s supply-side problem – the intermediary can now avail itself of additional supply by entering into a contract with a supplier (another water intermediary) or by purchasing directly from the spot market (if there is in fact a spot market from which to purchase). More generally, the presence of a market introduces an additional set of supply-side management tools.

Water bought by an urban water district (our intermediary of interest) from an agricultural water district represents a *reallocation* of water from an agricultural to an urban user. Reallocation of water through markets has been of interest in parts of the world where there are defined rights to water use. These water rights are often concentrated in the hands of agricultural users yet urban use is increasingly the higher-value use (or at least higher-cost supply). Prices for urban water have increased significantly over time (with increasing demand) to a point where they often eclipse the profits that a farmer can make by applying the same water to his crops. This has prompted speculation that there could be significant gains from trade between agricultural water rights holders and urban water users.

Contractual supply arrangements akin to those between generators and distributors in the electricity market would reallocate the water from agricultural users to urban users. These

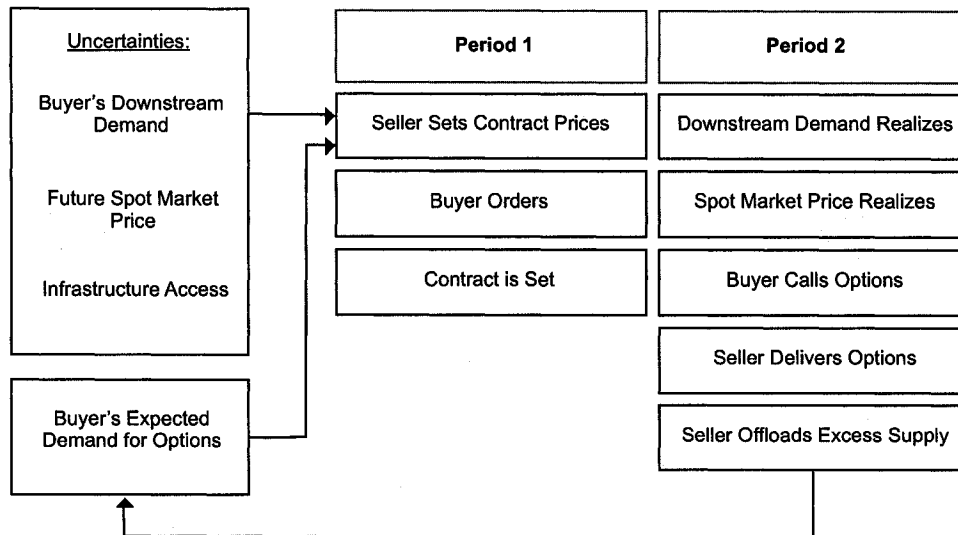


Figure 1.3: Bilateral Option Contracting

contracts would allow urban districts to avoid costly shortfalls in supply. In particular, option contracts would accomplish this aim while allowing water to remain in productive use (as an input to agricultural production) in times when existing urban supplies adequately cover demand.

In Chapter 3, we develop a bilateral option contracting model for the water market. The model draws from the experience in the electricity sector and, in particular, from the work on contracting in capital-intensive industries by Wu et al. (2002). We assume a two-period setting – in the second period, the buyer will decide how many options to exercise (out of the total that he purchased), given the demand realization. The seller will deliver those options and then offload his excess supply onto the spot market. In the first period, the seller sets prices (and the buyer purchases options). We solve the model in a perfect-information setting, first deriving the buyer's optimal purchase quantity, as a function of the prices set by the seller, and then solving the seller's first period problem, e.g., finding the optimal option and strike prices. We investigate the impact of changes in value in both the buyer's and the seller's reserve values on the optimal contract prices and quantities and the feasibility of contracting. We also compare the contracting outcomes to the first-best allocation. The discussion in Chapter 3 is based on Tomkins and Weber (2008).

## 1.6 Option Contract Competition

Chapter 4 extends the discussion in Chapter 3 to consider contracting in the presence of multiple sellers and buyers. The entrance of additional sellers into the marketplace changes the game significantly, as strategic considerations now impact the seller's pricing decision. Problems of strategic interaction between players generally fall in the realm of game theoretic modelling. One of the best-known examples of a simple game with a single variable (price) is that of Bertrand price competition. In the Bertrand model, two firms compete for customers, with both firms announcing price in the first period and then serving demand in the second period. The strong and perhaps surprising conclusion in this simplified setting is that with just two sellers, the competitive outcome is achieved. That is, the unique Nash equilibrium of the game is that where the two firms announce the same price, equal to their marginal cost, and make zero profits. As we discuss in Chapter 4, there are numerous wrinkles that, when introduced, relax this conclusion and arguably lead to more realistic predictions. For instance, both the introduction of asymmetric costs and capacity constraints lead to equilibria where the firms make a profit. Price competition between firms with capacity constraints is the case of interest to us in Chapter 4. Osborne and Pitchik (1984) study price competition between two sellers in the presence of capacity constraints and complete information. They identify pure-strategy equilibria in two capacity states: under-capacity and over-capacity. When capacity is in a mid-range the equilibria are in mixed strategies. We consider capacity-constrained price competition between two-sellers in the option contracting model in both a complete and incomplete information setting.

We conclude Chapter 4 with a brief discussion of contracting in a multi-agent ( $N > 2$ ) single-principal model (where here the sellers are the agents, bidding to serve the principal's, or buyer's, demand). The most familiar contracting model with multiple agents and a single principal is the auction model. We consider a multi-unit auction, implemented by the buyer. A full analysis of the auction problem is beyond the scope of this thesis. We instead focus our discussion on two aspects of the auction – efficiency and the distribution of surpluses – providing a numerical example as illustration.

## 1.7 Application to California Water

For an application of the models developed in Chapter 2 and Chapter 4, we turn to California. The California water system is of potential interest for several reasons. First, the current system is under significant pressure, and there is a generally recognized need for reform.

Second, there has been long-standing interest in the development of a water market in California to facilitate the reallocation of water from agricultural users to urban users. The disparate prices paid for water by agricultural and urban water users suggests that there can be significant gains from trade. The cost of urban water supply is a minimum of \$300/acre-foot (af) everywhere in Southern California, with prices often much higher. The average price per-af paid by agricultural districts in California is less than \$20/af. As discussed below, a number of market frictions have impeded the development of the water market in the state. Temporary water transfers, as achieved through option contracting, have the potential to reduce these frictions and stimulate development of the state's water market. Lastly, the challenges faced in California are representative of those that are being faced, or will be faced, elsewhere in the globe. As such, the solutions implemented in California are likely to be of international interest.

The state faces a number of challenges with respect to water allocation. Already the most populous state in the U.S., the population is projected to grow by almost 50% to reach over 45 million residents by 2020. This translates into increasing demand for water (largely in urban areas). The state's current urban water supply is already insufficient to meet demand during extended periods of drought, as evidenced in the 1986-1992 drought when storage supplies ran low and water was rationed to Southern California residents. Existing urban water supplies are, at the same time, becoming less reliable for several reasons: first, groundwater mining has resulted in a number of over-drafted aquifers – aquifers where pumping is either no longer feasible or will result in additional, undesirable, land subsidence. Second, the San Francisco Bay Delta, which serves as the major conduit for water from the northern to the southern part of the state, faces its own set of problems. The populations of several native fish species are in serious decline and listed as endangered under the federal Endangered Species Act (ESA). The ecosystem overall is in a state of decline. The levee system that is currently in place is inherently unstable as a result of its construction on unstable soils. If, or rather when, the system collapses, the Delta will no longer be a viable conduit of state water supplies. Both environmental concerns and water supply reliability concerns suggest a course of action that will ultimately bypass the Delta for water supply (and possibly translate into reduced supply moving from north to south in the short-term).

Third, a number of climate change impacts may reduce available water supply. Increases in temperature of several degrees centigrade, as predicted for the state, will result in higher evapotranspiration rates. Shifts in rainfall and runoff patterns and, in particular, in the timing of the snowpack melt may lead to lower storage volumes. The snowpack in the Sierras now melts earlier in the season due to global warming, and, hence, less water arrives



later in the year, when it is needed. Whether the state will receive more or less water on average under the new climate regime is unclear from current climate models, which predict an increase in precipitation in the Pacific Northwest and a decrease in the arid Southwest and a minor shift of indeterminate sign for California. Finally, salt water intrusion into coastal aquifers and the San Francisco Bay-Delta, due to rising sea levels, threatens freshwater supply.

Strategies for addressing water supply reliability issues in the face of climate change include reliance on additional above surface and subsurface storage, investment in additional water treatment facilities, including reuse and desalination plants, and reoperation of existing reservoir systems. The latter may, for instance, allow the system to adjust to the shift in timing of the snowpack melt. Reservoirs are operated to both store maximal amounts of water for use in dry periods and to provide flood control. Current operating practice is to draw down reservoir levels prior to the big spring snowpack melt and then capture and retain much of the meltwater. If the snowpack melt occurs late in the season, after the big rainstorm season in the winter and early spring, limited reservoir capacity has to be reserved for flood management. If, however, the melt shifts earlier, then the competing interests of flood protection and meltwater storage require reoptimization of the reservoir operations. Apart from investment in additional storage and reoperation of existing facilities, the state's water intermediaries may also look to reallocation as an avenue for addressing water supply reliability, as could be achieved voluntarily through water markets.

The potential of water markets to alleviate the state's water supply issues has captured attention but has thus far gone unrealized. The predominant focus of the academic discussion and the deal-making to date has been on the transfer of permanent water rights, also referred to as statutory transfers. These transfers have proved difficult to implement for several reasons. First, statutory transfers meet with institutional resistance. Farmers and members of agricultural communities recognize that selling water rights threatens their livelihoods and way of life. Even when a water rights sale might be desirable from the communities perspective, the charters of the irrigation districts (special districts formed under state legislative statute) don't specify a procedure for distributing proceeds from a water sale to the districts' members. The lack of an accepted rule for distributing profits from water sales often results in disagreements or simply makes district members' reluctant to approve transfers, having no guaranteed profit stake (Thompson 1993).

That the charters of state irrigation districts don't specify how to handle water transfers was an inevitable oversight when the districts were first formed in the 1900s, as they were designed to band together local water users and not to facilitate water transfers. This omission in the

district's charter leads to disagreements regarding the distribution of proceeds from water sales and can ultimately prevent transfers. Finally, specific members of districts may seek to block transfers and thereby keep the price of additional irrigation water for their own crops (artificially) low. For both legitimate and illegitimate reasons, institutions today block statutory water transfers.

Additional impediments to statutory transfers include the cost and time, and final approval, required to comply with the California environmental review process. There is also uncertainty regarding infrastructure availability to transfer water, which increases the uncertainty surrounding the ultimate value of the water obtained under transfer and can make agreement regarding a final price difficult. The contracting market in California and elsewhere depends critically on the both infrastructure integrity and availability.<sup>5</sup>

Temporary transfers reduce institutional resistance significantly by removing the threat to the agricultural communities' viability. They also require less intensive environmental review due to their temporary nature. In an option contract, the financial exposure related to infrastructure failure is significantly reduced (as the loss is reduced to the upfront fee if the water is ultimately untransferable). There are currently several examples of option contracting in the California water market. The Metropolitan Water District (MWD), the largest intermediary in the state, has entered into option contracts with Sacramento Valley farmers, in 2003, 2005, and again this year. This year another major water intermediary in the state, the San Diego County Water District (SDCWA), also signed option contracts, with two northern irrigation districts, for the first time.

The volume of transactions in the California water market ultimately depends on the availability of conveyance infrastructure. There are six major water arteries in the state. The State Water Project aqueduct conveys water from northern to Southern California. As discussed, the major bottleneck in this conveyance system is the Delta and, specifically, the operation of the Delta pumping facilities. If plans currently on the table to construct a peripheral canal and bypass the Delta go forward, the new infrastructure could facilitate a much greater volume of transfers between northern and Southern California.<sup>6</sup> There is currently excess conveyance capacity in the infrastructure south of the Delta. However, the volume of water available for interdistrict transfers is also more limited. For the water market

---

<sup>5</sup>Physical infrastructure such as canals and aqueducts clearly facilitates transfers. However, even without such infrastructure, water can be transported (expensively). In the developing world, for example, municipal water supply is not uncommonly supplemented by tanker delivery.

<sup>6</sup>This is notably one of the major reasons for political opposition to the peripheral canal: a desire of residents in the northern part of the state to prevent additional water appropriation by Southern Californians.

to reach its reallocation potential will likely require major infrastructure investments. It may also require restructuring of infrastructure control to ensure equal access by participating intermediaries.

# Chapter 2

## Joint Inventory-Price Control

### 2.1 Motivation

Interest in inventory control and joint inventory-price control models is longstanding, dating back to pioneering work in the early 1950s by Arrow et al. (1951) and Whitin (1953, 1955). The traditional application of such models has been to the retail sector: a retailer faces the joint decision of how many items to order and what price to charge, given his ordering cost, the holding cost associated with storing inventory, the shortage cost associated with unmet demand, and a (possibly uncertain) demand function. The joint inventory-price control model provides a natural framework for addressing a water intermediary's short-run allocation problem. The water intermediary is interested in establishing an optimal ordering (water procurement) and pricing policy in the presence of demand (and supply) uncertainty, with the option to store water.

In this chapter, we develop an inventory and price control model for application to a large water intermediary in the public sector. In doing so, we extend the current model to treat both profit maximization and welfare maximization, where the latter is presumed to be of primary interest to an intermediary operating in the public sector. We dispense with the common "backlogging assumption," in which all lost demand is backlogged and then filled in the next period. This assumption is unrealistic for water deliveries, where at least demand for municipal and industrial (M&I) water is of an immediate nature and cannot be filled at a later date.<sup>1</sup> We also introduce supply uncertainty to the model. We make several additional

---

<sup>1</sup>The large water intermediary studied herein, the Metropolitan Water District of Southern California, supplies several different classes of water: 'full service treated' water, which is generally for M&I use, 'interim

modifications to the standard model, as follows: (1) we specify an explicit capacity level for inventory storage, (2) we introduce a variable termed *contract size*, which limits the order level, and (3) we define a loss coefficient to capture possible transmission losses in the system.

As with previous expositions on inventory and pricing control models, we will be interested in characterizing the optimal solution to our model. Under the backlogging assumption, Federgruen and Heching (1999) established that the optimal policy in a joint inventory-price control model with pure demand uncertainty is a base-stock-list-price policy, or  $(s, S, P)$  policy. That is, there is a base-stock level  $s$  such that whenever the existing inventory level falls below  $s$ , an order is placed to bring the current inventory level up to  $S$ . If the current inventory state is greater than  $s$ , no order is placed; the list price is that price  $P$  which is optimal for  $S$ . The base-stock-list-price policy dictates an optimal policy when the constraint that the order-up-to level, or base stock, be less than the current inventory state is non-binding. In the case that the constraint is binding, it is optimal not to order any additional items, but the policy does not explicitly dictate the optimal price. Federgruen and Heching (1999) establish that the price is non-increasing in the inventory state, which in turn implies the intuitively appealing policy of price discounting in the case of excess inventory.

The optimality of this base-stock-list-price policy hinges on the concavity of the revenue function, which in turn ensures concavity of the payoff function. Without the backlogging assumption, we are assured neither concavity of the revenue function nor, ultimately, of the payoff function. We show that the optimal solution in the relaxed inventory and pricing control model can be characterized generally as one in which the optimal order level is increasing in both the current inventory state and the contract size. However, somewhat surprisingly, the optimal price can be either decreasing *or* increasing in the current inventory state. Specifically, the single-period payoff function that Federgruen and Heching (1999) showed to be submodular in the inventory order level and price pair  $(y, p)$  may become supermodular (for instances of highly inelastic demand). The optimal price  $p^*$  is decreasing (increasing) in the current inventory state  $x$  for a submodular (supermodular) payoff function. This result holds only in the case of profit maximization and not for the case of welfare maximization. The implication of the result is that a profit-maximizing intermediary facing highly inelastic

---

agricultural' water, and 'storage replenishment' water. The latter might be more accurately modelled under the assumption that backlogging is permitted, since if a storage facility goes unfilled in this period it will still need replenishment in the next period. Demand for storage water is effectively carried over to the next period. However, when modelling M&I supplies (our primary interest), the no backlogging assumption that we adopt is appropriate. To the extent that M&I water is being pre-ordered and stored, however, the assumption of at least partial backlogging may be even more reasonable.

demand will choose not to lower (and may even raise) prices in periods of “excess” inventory. The introduction of supply uncertainty does not materially change the model structure. It does, however, introduce two new parameters – the variance of supply and the correlation between supply and demand. In our discussion of monotone comparative statics of the model, we show that for normally distributed (and positively correlated) supply and demand shocks, the optimal inventory order levels and prices are increasing in the mean of the shocks. That is, a shift in the distributions of demand and/or supply to make higher demand states and/or lower supply states more likely results in an increase in the optimal inventory order level and optimal price. More generally, for a given density (of the joint supply and demand uncertainty), the optimal inventory-price pair is increasing in any parameter for which the density satisfies the monotone likelihood ratio (MLR), e.g., the mean of the bivariate normal density.

Implications of the model for a water intermediary are explored in Section 2.4. The application of the model is to the Metropolitan Water District of Southern California (MWD), the largest water intermediary in California, with a current operating budget of over a billion dollars. MWD is a major player in the quasi-monopolistic water distribution in the state. As such, it contracts with the State Water Project (SWP) for the majority of its water (over two-thirds, or 2 million acre feet (maf) annually). Contracts for Colorado River water and various options contracts and fallowing agreements with farmers, as well as conjunctive use agreements, constitute the remainder of MWD’s supply. In addition, MWD manages a number of reservoirs and groundwater storage programs.

MWD’s demand and supply are subject to significant random shocks. On the supply side, the available upstream water resources are largely determined by the precipitation levels and the hydrological conditions in Northern California.<sup>2</sup> In general, both municipal and agricultural water demand increases with temperature.

MWD primarily supplies water to its member agencies. There are currently 26 member agencies – mostly city or county water agencies or authorities, including the Los Angeles Department of Water and Power and the San Diego County Water Authority. The member agencies are voting members of MWD’s board, with voting rights proportional to the assessed valuation of the land within the agency’s service area. As such, they review the annual rates proposed by MWD’s staff. The current rate structure is designed to recuperate average costs.<sup>3</sup> A fixed fee is charged to each member agency, billed as a “Capacity Charge” and

---

<sup>2</sup>As MWD is contracting with SWP and not directly diverting water from the source, its supply, as we shall see, is not as closely coupled to the hydrologic cycle and is, rather, affected by the SWP operations.

<sup>3</sup>There is a “Water Rate Stabilization Fund,” which helps offset revenue shortages in years of lower

a “Readiness-to-Serve Charge,” and the remaining costs are recovered through a per-unit charge per af of delivered water.

The rest of this chapter is organized as follows: in Section 2.2 we develop the inventory-price control model with two-sided uncertainty. In Section 2.3, we establish the intermediary’s optimal policy and discuss comparative statics of the model. Section 2.4 presents results from the numerical study of MWD. Finally, in Section 2.5 we provide a brief summary.

## 2.2 The Model

Consider a dynamic water distribution system, the state of which is characterized at each time  $t \in \{0, 1, \dots\}$  by the available water inventory  $x_t \in X = [0, \bar{x}]$ , where  $\bar{x} > 0$  corresponds to the total storage capacity. The water distribution system is run by an intermediary, who, at each time  $t$ , takes decisions about the conveyance price  $p_t \in \mathbb{R}$  and the order-up-to level  $y_t \in [x_t, x_t + z]$ , where  $z$  represents an upper bound on the quantity ordered at time  $t$ .<sup>4</sup> The length of a discrete time interval in the analysis is chosen in accord with the actual schedule of decisions.<sup>5</sup>

The demand  $\tilde{d}_t$  for water at time  $t$  is a random variable, the expectation of which is decreasing in the price  $p_t$  at time  $t$ . For simplicity, we consider here an affine dependence of (nonnegative) demand on price with additive uncertainty, so that

$$\tilde{d}_t(p_t) = [\alpha - \beta p_t + \tilde{\varepsilon}_t]_+,$$

where  $\alpha, \beta$  are positive real constants and  $\tilde{\varepsilon}_0, \tilde{\varepsilon}_1, \dots \sim N(0, \sigma_\varepsilon^2)$  are i.i.d. zero-mean normal random variables with variance  $\sigma_\varepsilon^2$ . The water supply  $\tilde{s}_t$  available at time  $t$  to satisfy demand is a random variable, the realization of which depends on the order-up-to level  $y_t$ ,

$$\tilde{s}_t(y_t) = \left[ \min \left\{ \lambda y_t + \tilde{\xi}_t, \bar{x} \right\} \right]_+,$$

where  $\lambda \in (0, 1]$  describes a possible transmission loss and  $\tilde{\xi}_0, \tilde{\xi}_1, \dots \sim N(0, \sigma_\xi^2)$  are i.i.d. zero-mean normal random variables with variance  $\sigma_\xi^2$ . The available supply  $\tilde{s}_t$  cannot exceed the total storage capacity  $\bar{x}$ . The quantity of water delivered (or *demand served*) at time  $t$  is

$$\tilde{q}_t(p_t, y_t) = \min \left\{ \tilde{d}_t(p_t), \tilde{s}_t(y_t) \right\}.$$

---

deliveries, using the balance from high-demand years.

<sup>4</sup>This upper bound may, for instance, be specified in the contractual agreement(s) that the intermediary holds with its supplier(s). Alternately, it may represent the conveyance, or distribution, capacity.

<sup>5</sup>In Section 2.4, we discuss an application where decisions about order contracts and prices are made on an annual basis.

The cost  $C(x_t, y_t)$  of delivering water supplies in a given time period  $t$  depends on the (non-negative) difference between order-up-to level  $y_t$  and the currently held water inventory  $x_t$ :

$$C(x_t, y_t) = c(y_t - x_t)$$

where  $c$  denotes a positive unit delivery cost. The unit cost of storing, or holding, water inventory is  $h > 0$ , while the cost of unmet demand is  $r > 0$ , so that the expected loss associated with inventory deviations from demand at time  $t$  is

$$L(p_t, y_t) = E \left[ \max \left\{ h \left( \tilde{s}_t(y_t) - \tilde{d}_t(p_t) \right), r \left( \tilde{d}_t(p_t) - \tilde{s}_t(y_t) \right) \right\} \middle| p_t, y_t \right].$$

In inventory-control models with backlogged demand, the cost of unmet demand,  $r$ , is usually measured as the cost of fulfilling backlogged orders. In our model,  $r$  has a different interpretation, in keeping with the public pricing and, specifically, peak-load-pricing literature. Here,  $r$  is the shortage cost, which is defined in the peak-load-pricing literature as comprised of two parts: the welfare loss associated with curtailed supply and the disruption (or outage) cost, where the latter includes the cost of spoiled goods and lost productivity (Crew et al. 1995).<sup>6</sup> The welfare loss is already accounted for in our model, since we are modelling *demand served* rather than actual demand. Hence, the shortage cost  $r$  includes just the outage costs, i.e., loss in productivity and the cost of spoiled goods.

We treat both profit maximization and welfare maximization in the model. With regards to the latter, the expected consumer surplus associated with the linear demand function is given as

$$CS(p_t, y_t) = E[(\hat{p}_t - p_t)\tilde{q}(p_t, y_t) + (\tilde{q}(p_t, y_t))^2/(2\beta)]$$

where  $\hat{p}_t(\tilde{q}_t, \tilde{\varepsilon}_t) = (\alpha + \tilde{\varepsilon}_t - \tilde{q}_t)/\beta$  corresponds to a fictitious (random) market clearing price at the (random) quantity  $\tilde{q}_t$ . This clearing price can never be smaller than the actual price  $p_t = (\alpha + \tilde{\varepsilon}_t - \tilde{d}_t)/\beta$ , since by definition the demand served cannot exceed the actual demand, whence  $CS(p_t, y_t) \geq E[\tilde{q}^2(p_t, y_t)|p_t, y_t]/(2\beta)$ . Let

$$Q_t(p_t, y_t) = E[\tilde{q}_t(p_t, y_t)|p_t, y_t]$$

denote the expected demand served at time  $t$ . Given a discount factor  $\delta \in (0, 1)$ , and a policy  $\mu = \{\mu_0, \mu_1, \dots, \mu_T\}$  with  $\mu_t(x_t) = (p_t(x_t), y_t(x_t))$  the intermediary's total discounted expected payoff at time  $t \in \{0, \dots, T-1\}$  is recursively defined as follows:

$$\Pi_T(x_T; \mu) = p_T Q_T(p_T, y_T) - c(y_T - x_T) - L(p_T, y_T) + \omega CS(p_T, y_T),$$

---

<sup>6</sup>In an electricity outage, the cost of spoiled goods, for example, may include perishable goods, whereas in a period of water shortage it may include perennial crops such as tree crops and/or landscape plants.



for  $t = T$ , where

$$\omega = \begin{cases} 1, & \text{if the intermediary is welfare-maximizing,} \\ 0, & \text{otherwise,} \end{cases}$$

and

$$\Pi_t(x_t; \mu) = [p_t Q_t(p_t, y_t) - c(y_t - x_t) - L(p_t, y_t) + \omega \text{CS}(p_t, y_t) + \delta E[\Pi_{t+1}(\tilde{x}_{t+1}; \mu) | p_t, y_t]]_{(p_t, y_t) = \mu(x_t)}$$

for  $t \in \{0, \dots, T-1\}$ , where the stochastic state transition from any  $x_t$  to a (random) state  $\tilde{x}_{t+1}$  is defined by

$$\tilde{x}_{t+1} = \max \left\{ 0, \tilde{s}_t(y_t) - \tilde{d}_t(p_t) \right\}.$$

Optimizing over all admissible policies, the associated discrete-time dynamic-programming equation is

$$\Pi_t(x_t) = cx_t + \max_{(p_t, y_t) \in \mathbb{R} \times [x_t, x_t+z]} \{p_t Q_t(p_t, y_t) - cy_t - L(p_t, y_t) + \omega \text{CS}(p_t, y_t) + \delta E[\Pi_{t+1}(\tilde{x}_{t+1}) | p_t, y_t]\},$$

for all  $t \in \{0, \dots, T-1\}$ , and

$$\Pi_T(x_T) = cx_T + \max_{(p_T, y_T) \in \mathbb{R} \times [x_T, x_T+z]} \{p_T Q_T(p_T, y_T) - cy_T - L(p_T, y_T) + \omega \text{CS}(p_T, y_T)\}$$

for  $t = T$ .

The distribution of transition probabilities from any inventory state  $x_t$  to the inventory state  $x_{t+1} = x \in [0, \bar{x}]$ , as a function of choice variables  $(p_t, y_t)$ , is

$$F(x | p_t, y_t) = P(\tilde{x}_{t+1} \leq x | p_t, y_t) = \int_{-\infty}^{\infty} \left( \int_{\min\{\lambda y_t + \xi, \bar{x}\} - (\alpha - \beta p_t) - x}^{\infty} \varphi(\varepsilon, \xi) d\varepsilon \right) d\xi,$$

where  $\varphi$  denotes the joint density for the possibly correlated demand and supply uncertainties. For concreteness, we consider in this model a bivariate normal distribution with standard deviations  $\sigma_\varepsilon, \sigma_\xi > 0$  and correlation  $\rho \in (-1, 1)$ ,

$$\varphi(\varepsilon, \xi) = \frac{\exp \left[ -\frac{1}{2(1-\rho^2)} \left( \frac{\varepsilon^2}{\sigma_\varepsilon^2} - \frac{2\rho\varepsilon\xi}{\sigma_\varepsilon\sigma_\xi} + \frac{\xi^2}{\sigma_\xi^2} \right) \right]}{2\pi\sigma_\varepsilon\sigma_\xi\sqrt{1-\rho^2}}.$$

For  $x = \bar{x}$  we obtain  $F(\bar{x} | p_t, y_t) = 1$ , so that  $F$  may be discontinuous at the upper bound. Correspondingly, we find for the probability density for the state  $x_{t+1}$  that

$$f(x | p_t, y_t) = \int_{-\infty}^{\infty} \varphi(\min\{\lambda y_t + \xi, \bar{x}\} - (\alpha - \beta p_t) - x, \xi) d\xi$$

for all  $x \in (0, \bar{x})$  with atoms at the interval ends of size  $F(0 | p_t, y_t)$  and  $1 - F(\bar{x} | p_t, y_t)$  respectively. We can now rewrite the expected next-period payoff in the above dynamic-programming equation in the form

$$E[\Pi_{t+1}(\tilde{x}_{t+1}) | p_t, y_t] = \int_0^{\bar{x}} \Pi_{t+1}(x) f(x | p_t, y_t) dx + F(0 | p_t, y_t) \Pi_{t+1}(0) + (1 - F(\bar{x} | p_t, y_t)) \Pi_{t+1}(\bar{x}),$$

for all  $t \in \{0, \dots, T-1\}$ .

REMARK 1 When discretizing the state space to  $N + 1$  inventory levels  $\hat{x}^k = k\bar{x}/N$  for  $k \in \{0, \dots, N\}$ , the corresponding transition probabilities from  $x_t = \hat{x}^j$  to  $x_{t+1} = \hat{x}^k$  are given approximately by

$$\pi_{j,k}(p_t, y_t) = F((2k + 1)\bar{x}/(2N)|p_t, y_t) - F((2k - 1)\bar{x}/(2N)|p_t, y_t)$$

for  $k \in \{1, \dots, N - 1\}$ ,  $\pi_{j,N}(p_t, y_t) = 1 - F((2N - 1)\bar{x}/(2N)|p_t, y_t)$ , and  $\pi_{j,0}(p_t, y_t) = F(\bar{x}/(2N)|p_t, y_t)$ . Note that all the transition probabilities are independent of the state  $x_t$  at time  $t$ , i.e., independent of  $j$ , which simplifies computations substantially. In the discretized state-variables, the expected value in the dynamic-programming equation becomes (for all  $t \in \{0, \dots, T - 1\}$  and all  $k \in \{0, \dots, N\}$ )

$$E[\Pi_{t+1}(\tilde{x}_{t+1})|p_t, y_t] = \sum_{k=0}^N \pi_k(p_t, y_t) \Pi_{t+1}(k\bar{x}/N),$$

where, for simplicity, we have written  $\pi_k$  instead of  $\pi_{j,k}$ . □

## 2.3 Optimal Policy and Comparative Statics

In this section we characterize the optimal dynamic inventory-price policy and discuss the comparative statics of the model. For this it is useful to start by considering the comparative statics of the intermediary's expected per-period payoff, which we do below.

### Price and Order-Level Dynamics

The intermediary's expected per-period payoff is given as

$$g(p, y) + cx = pQ(p, y) - L(p, y) - c(y - x) + \omega CS(p, y),$$

where, for convenience, we drop the dependence of variables on the time  $t$  unless there can be misunderstandings.

**Proposition 2.1** *The expected demand served  $Q(p, y)$  and the expected consumer surplus  $CS(p, y)$  are both decreasing in  $p$ , increasing in  $y$ , and submodular in  $(p, y)$ .*

As water is a normal good, a higher price is unambiguously bad news for consumers, hence both expected demand-served and consumer surplus are necessarily decreasing in price. The proof of Proposition 2.1 shows that the marginal decrease of expected demand served  $Q$  with

respect to price is proportional to the demand-elasticity parameter  $\beta$  and the probability that supply exceeds demand,

$$Q_p(p, y) = -\beta P(\tilde{d}_t(p) \leq \tilde{s}_t(y)) \leq 0.$$

Conversely, the marginal increase of  $Q$  with respect to the order-up-to level  $y$  is proportional to the loss factor  $\lambda$  and the probability that supply neither reaches capacity nor fully covers demand,

$$Q_y(p, y) = \lambda P(\tilde{s}_t(y) < \min\{\tilde{d}_t(p), \bar{x}\}) \geq 0.$$

Moreover, the submodularity of  $Q$  implies that  $Q_p$  decreases with an increase in  $y$  and, vice versa,  $Q_y$  decreases with an increase in  $p$ . A submodular function can also be characterized as one experiencing “decreasing differences” (also referred to as antitone differences in the literature). Decreasing differences can be interpreted as follows: in the presence of a higher level of  $y$  it is desirable to have a *lower* level of  $p$ , and vice versa. Increasing differences underlies the economic notion of complementarity - that when you have more of good  $a$  it is better to also have more of good  $b$ . As we discuss in Section 2.3.2 these properties play a central role in the comparative statics of the model.

The marginal decrease of expected consumer surplus  $CS$  with respect to a unit price increase is proportional to the expected demand served,

$$CS_p(p, y) = -Q(p, y) \leq 0.$$

The marginal increase of  $CS$  with respect to a unit increase in the order-up-to level  $y$  is proportional to the expected excess demand, conditional on the fact that supply neither reaches capacity nor fully covers demand,

$$CS_y(p, y) = \frac{\lambda}{\beta} E \left[ \tilde{d} - \tilde{s} \mid \tilde{s} < \min\{\tilde{d}, \bar{x}\} \right] P(\tilde{s} < \min\{\tilde{d}, \bar{x}\}) \geq 0.$$

The nonnegative expected losses  $L(p, y)$  are not monotonic with respect to either price or order-up-to level. For any given positive holding and shortage cost parameters  $h$  and  $r$ , they vanish if and only if demand is exactly met by supply. The marginal losses with respect to a unit change in price can be determined analogously to the computations in the proof of Proposition 2.1:

$$L_p(p, y) = -\beta r + \beta(r + h)P(\tilde{d}_t(p) \leq \tilde{s}_t(y)),$$

$$L_y(p, y) = -\lambda r + \lambda(h + r)P(\tilde{d}_t(p) \leq \tilde{s}_t(y)),$$

and

$$L_{py}(p, y) = \beta\lambda(r + h) \int_{-\infty}^{\bar{x} - \lambda y} \varphi(\lambda y + \xi - (\alpha - \beta p), \xi) d\xi \geq 0.$$

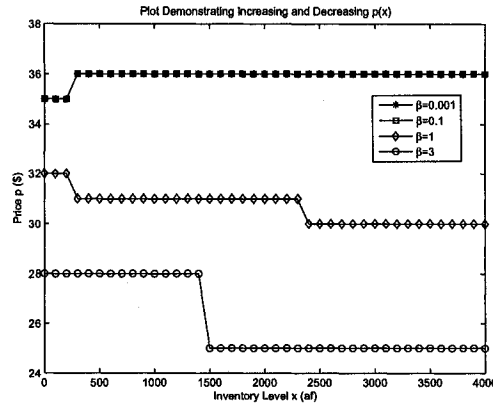


Figure 2.1: A Price Increase Under Supermodularity

Hence, the term  $-L(p, y)$  in the intermediary's per-period payoff is submodular with respect to  $(p, y)$ . Using the results above, we can examine the sub- or supermodularity of the per-period payoff function,  $g(p, y)$ , which, in turn, will be useful for characterizing the price dynamics in the one-period model. We begin with Proposition 2.2 below.

**Proposition 2.2** (i) Under welfare maximization ( $\omega = 1$ ), the intermediary's per-period payoff is submodular in  $(p, y)$ . (ii) Under profit maximization ( $\omega = 0$ ), the intermediary's per-period payoff may become supermodular in  $(p, y)$  provided that demand is sufficiently inelastic, e.g., for small enough  $\beta$ .

The submodularity and supermodularity results come from examination of  $g_{py}(p, y)$ , where, in the case of welfare maximization,  $g_{py}(p, y)$  is seen to be always negative. The negative impact of an increase in price - which decreases expected demand served and, hence, decreases expected consumer surplus as well as expected revenue - and the cost to the intermediary of ordering and storing an additional unit of inventory outweighs the expected gain in revenue and consumer surplus from an increase in  $y$ . However, under profit maximization when the possibility that the demand is (largely) unaffected by a price increase exists, i.e., for  $\beta \ll 1$ ,  $g_{py}(p, y) > 0$ . The sub- and supermodularity results imply the following Proposition.

**Proposition 2.3** Let  $t \in \{0, \dots, T\}$ . (i) Under welfare maximization ( $\omega = 1$ ), the optimal price  $p_t^*(x)$  is nonincreasing in  $x$ . (ii) Under profit maximization ( $\omega = 0$ ), the optimal price  $p_t^*(x)$  is generically nonmonotonic; in particular it can be locally increasing in  $x$ .

Part (ii) of the last result is somewhat counterintuitive and contrasts with results for the standard model (see e.g., Federgruen and Heching (1999)). Consider the last period  $t = T$  in which the intermediary cares only about the per-period payoff  $g(p, x) + cx$ , provided that  $y^*(x) = x$ . When demand is extremely inelastic (i.e., for small  $\beta$ ), then the positive marginal benefit of a price increase for the intermediary's payoff increases with the inventory level, since that price increase can be paid by a larger number of consumers. Figure 2.3 shows a numerical example, in which – to obtain the effect at a reasonable price – the optimal price was held close to a target level  $p_{\text{target}}$  by adding a negative quadratic term  $-k(p - p_{\text{target}})^2$ , for some  $k > 0$ , to the intermediary's per-period payoffs, which penalizes deviations from the target price, and does not influence the supermodularity properties of  $g$ . For the example,  $p_{\text{target}} = 20$ ,  $k = 6$ , and the values of the other parameters are given as follows:  $\alpha = 174$ ,  $\beta = 3$ ,  $c = 22$ ,  $h = 0.20$ ,  $r = 22$ , and  $\bar{x} = 400$ . For  $\beta \leq 0.1$ , the optimal price  $p^*(x)$  increases from \$35 to \$36 as the inventory level  $x$  increases from 0 to 500 af.

At any time  $t$ , both a higher price and a higher order-up-to level tends to increase the stochastic next-period inventory level. The decision vector  $(p_t, y_t)$  induces a first-order stochastic dominance (FOSD) order of the distributions of  $\tilde{x}_{t+1}$ , as established in Proposition 2.4

**Proposition 2.4** *The distributions  $F(\cdot|p, y)$  corresponding to the state-transition probabilities conditional on  $(p, y)$  are FOSD-ordered in the sense that*

$$(p, y) \leq (\hat{p}, \hat{y}) \quad \Rightarrow \quad F(\cdot|p, y) \preceq_{\text{FOSD}} F(\cdot|\hat{p}, \hat{y})$$

for all  $(p, y), (\hat{p}, \hat{y}) \in \mathbb{R} \times X$ .

### 2.3.1 Optimal Policy

The standard form of the optimal inventory-price control policy, established by Federgruen and Heching (1999), is the base-stock-list-price, or  $(s, S, P)$  policy in which the intermediary orders up to level  $S$  whenever inventory is below level  $s$  and sets list price  $P$ . This result holds under the backlogging assumption (preserving concavity of the revenue function and, hence, joint concavity of the payoff function in  $p$  and  $y$ ). We have relaxed the backlogging assumption and are thus not ensured concavity of the revenue function. When joint concavity holds, the optimal policy under two-sided uncertainty is the base-stock-list-price policy. In general, we show the optimal order level to be increasing in both the current inventory level  $x$  and the contract size  $z$ .

**REMARK 2** The intermediary's expected gross revenues  $pQ(p, y)$  are generally not concave

in  $(p, y)$ , even though the expected demand served is. To see the latter, note that

$$\begin{aligned} Q_{pp}(p, y) &= -\beta^2 \int_{-\infty}^{\bar{x}-\lambda y} \varphi(\lambda y + \xi - (\alpha - \beta p), \xi) d\xi < 0, \\ Q_{py}(p, y) &= -\beta\lambda \int_{-\infty}^{\bar{x}-\lambda y} \varphi(\lambda y + \xi - (\alpha - \beta p), \xi) d\xi < 0, \\ Q_{yy}(p, y) &= -\lambda^2 \int_{-\infty}^{\bar{x}-\lambda y} \varphi(\lambda y + \xi - (\alpha - \beta p), \xi) d\xi < 0. \end{aligned}$$

Let  $H(Q) = \partial^2 Q(p, y) / \partial(p, y)^2$  be the Hessian matrix for  $Q$ . Then  $\det H(Q) = Q_{pp}Q_{yy} - Q_{py}^2 \leq 0$  and  $\text{trace } H(Q) = Q_{pp} + Q_{yy} < 0$ , so that both eigenvalues of  $H(Q)$  must be negative. This implies that  $H(Q)$  is negative definite and thus  $Q$  strictly concave in  $(p, y)$ . However, it can be shown using an analogous method that the function  $pQ(p, y)$  may not be concave in  $(p, y)$ .  $\square$

We begin by establishing the existence of a nonzero, finite optimal price. The idea is to show that for any fixed order-up-to level  $y$  the intermediary's payoffs are 'coercive' in price, in the sense that it is always better for the intermediary to charge a price larger than zero and less than infinity, implying the existence of an optimal finite price.

**Proposition 2.5** *For any  $t \in \{0, \dots, T\}$  and any  $x \in [0, \bar{x}]$  the optimal price  $p_t^*(x)$  lies in  $[0, \bar{p}(x)]$ , where  $\hat{p}(x) = \inf \left\{ p \geq 0 : P(\tilde{d}_t(p) \leq \tilde{s}_t(x)) \geq 1/2 \right\}$  and  $\bar{p}(x) = \max \left\{ \hat{p}(x), r + h - \omega \left( \frac{\alpha}{\beta} - \frac{2\zeta}{\beta} \right) + \rho + \frac{2\bar{x}}{\beta} \right\}$ , where*

$$\zeta = \left( E[\tilde{\varepsilon} | \tilde{d}_t(p) < \tilde{s}_t(y)] - E[\tilde{q} | \tilde{d}_t(p) < \tilde{s}_t(y)] \right) P(\tilde{d}_t(p) < \tilde{s}_t(y)).$$

*The upper bound  $\bar{p}(x)$  is nonincreasing in  $x$ .*

The result below provides an explicit upper bound for the optimal price, corresponding to a price level at which the reduction in demand outweighs all potential benefits of a profit-maximizing intermediary, where

$$u_t(p_t, y_t; x_t) = cx_t + p_t Q_t(p_t, y_t) - cy_t - L(p_t, y_t) + \omega \text{CS}(p_t, y_t) + \delta E[u_{t+1}(\tilde{x}_{t+1}) | p_t, y_t] \quad (2.1)$$

is the intermediary's payoff under  $(p_t, y_t)$ . It is clear that because of  $\text{CS}_p < 0$  a welfare-maximizing intermediary never charges a higher price than a profit-maximizing intermediary. As the available inventory  $x$  increases, the upper bound  $\bar{p}(x)$  for the price weakly decreases.

**Proposition 2.6** (i) For any  $t \in \{0, \dots, T\}$ ,  $\arg \max_{(p_t, y_t) \in \mathbb{R} \times [x_t, x_t + z]} u_t(p_t, y_t; x_t)$  is increasing in the inventory level  $x$  and increasing in the contract size  $z$ . For submodular (supermodular)  $\Pi_t$ ,  $p^*$  is decreasing (increasing) in  $x$  and  $z$ . (ii) If the revenue function  $pQ(p, t)$  is concave in  $(p, y)$ , then the optimal policy can be characterized as a base-stock-list-price policy, or  $(s, S, P)$  policy.

That the optimal solution is increasing in the starting inventory level  $x$  can be established using the following argument. For a given  $x$ , we are looking for a solution  $y$  in the window  $[x, x + z]$ . Let  $y^*$  denote the optimal order level for this window. Now consider  $[x', x' + z]$ , where  $x' > x$ . Let  $y'^*$  denote the (new) optimal order level for this shifted window. Suppose that  $y'^* < y^*$ . But if  $y'^* < y^*$ , then  $y'^*$  was also a possible order level in the previous window  $[x, x + z]$ , which contradicts the optimality of  $y^*$ . We see that shifting the window to the right only creates the possibility of higher optimal order levels. A similar argument holds for the case of  $[x, x + z']$ , where  $z' > z$ .

(Note that the existence of an optimal solution is guaranteed by the continuity of  $\Pi_t$  and the compactness of the solution set for  $y$  and for  $p$ , where compactness for  $p$  was established in Proposition 2.5.)

The proof of part (ii) of Proposition 2.6, owing to Federgruen and Heching (1998), shows that the result in the operations management literature on optimal inventory-control and pricing policies under uncertainty can be extended to the case of two-sided uncertainty with lost sales (no demand backlogging) and order level  $y \in [x, x + z]$ , provided that the revenue function is concave. In which case, it is optimal for the intermediary to follow a base-stock-list-price policy, also termed  $(s, S, P)$ -policy, in which the intermediary orders up to a base-stock level  $y^*$  whenever the current water  $x$  is below  $y^*$  and does not order otherwise. The price  $p^*$  is the optimal price at  $y^*$ .

### 2.3.2 Comparative Statics

In Proposition 2.2 above, we establish submodularity of the payoff function under welfare-maximization. (Under profit-maximization, super- or submodularity of the payoff function depends on the magnitude of the model parameters, e.g., for small enough  $\beta$  the payoff function is actually supermodular.) For our discussion of comparative statics we will find it useful to work with an analogue of this result, given in Proposition 2.7 below. One well known comparative statics result (Milgrom and Shannon (1994)) states that if the payoff function exhibits increasing differences with respect to a parameter of interest and each of

the decision variables and the payoff function is supermodular *then* the optimal policy is increasing in the parameter.

**Proposition 2.7** (i) Under welfare maximization ( $\omega = 1$ ), the intermediary's per-period payoff is supermodular in  $(-p, y)$  and, equivalently, in  $(p, -y)$ . (ii) Under profit maximization ( $\omega = 0$ ), the supermodularity of the intermediary's per-period payoff function in  $(-p, y)$  (equivalently,  $(p, -y)$ ) depends on the specific choice of parameters.

The equivalence of supermodularity in  $(-p, y)$  and  $(p, -y)$  leads to incomplete comparative statics conclusions. For an increase in a given parameter (assuming increasing differences), we can conclude that the intermediary will adopt one of two policy shifts: he will either increase  $p$  and decrease  $y$  or he will decrease  $y$  and increase  $p$ . We cannot predict which course of action he will ultimately adopt. While this may seem unsatisfactory, we note that it is a rather natural result in the context of our inventory-price control problem, where price  $p$  and order level  $y$  can be thought of as strategic substitutes for the intermediary. For the comparative statics results discussed below, we assume that supermodularity in  $(-p, y)$  (equivalently,  $(p, -y)$ ) holds, i.e., assumption A1 below is in effect.

A1. The per-period payoff function is supermodular in  $(-p, y)$  (equivalently  $(p, -y)$ ).

REMARK 3 In referring to Proposition 2.7 as an analogue of Proposition 2.2 we do not mean to imply that the relationship between super- and submodularity of the control variables in our model is one that holds in general. It is not generally true that a function that is submodular in a given pair of decision variables  $(p, y)$  is then supermodular in  $(-p, y)$  and  $(p, -y)$ . For an easy counter-example, consider the function  $f(p, y) = -p^2y$ . This function is submodular in  $(p, y)$ , also submodular in  $(-p, y)$ , and supermodular in  $(p, -y)$ , as can easily be verified by calculating each of the cross-partials. In the joint inventory-price control model, the supermodularity in  $(-p, y)$  and  $(p, -y)$  highlights the particular interaction between the price and order-level, namely that the adjustment of either control will achieve the same directional impact. An increase in price will, in general, depress demand, reducing shortage costs. Similarly, an increase in the order level will reduce expected shortages. An intermediary would not raise price, thereby suppressing demand, and at the same time order more! Similarly, if the order level is increased, it would be counter-productive to then raise price and stymie demand.  $\square$



Our first comparative statics result establishes the intuitively appealing fact that a welfare-maximizing intermediary will charge a lower price  $p$  and set a higher order level  $y$  than its profit-maximizing counterpart. Proposition 2.8 establishes the fact that a social welfare maximizing intermediary will charge a lower price and set a higher order-level than a profit-maximizing intermediary.

**Proposition 2.8** *The arg max $_{(-p_t, y_t) \in \mathbb{R} \times [x_t, x_t + z]}$   $u_t(p_t, y_t; x_t)$  is increasing in  $\omega$ , where here we allow  $\omega$  to be any positive weight.*

The proof of the proposition relies on the observation that the derivative of the consumer surplus term is negative with respect to price and positive with respect to order level – an increase in price or a decrease in the order level are both bad news for consumers. The value function exhibits increasing differences in  $(-p, y)$  and  $\omega$ . This, coupled with the supermodularity of the value function under assumption A1, suggests that the optimal price is decreasing in  $\omega$  while the optimal order level is increasing.

Proposition 2.9 introduces a second intuitively appealing result: the payoff function is decreasing in all costs, including the shortage cost  $r$ , the storage, or holding, cost  $h$ , and the unit order cost  $c$ . The proof is straightforward. The value function is monotone decreasing in costs. By the linearity of the expectation operator, the single-period payoff function will be monotone decreasing in a given cost parameter, and, as the costs don't affect the stochastic transitions in the model, this result extends to the dynamic multi-period model.

**Proposition 2.9** *The discounted expected payoff function  $\Pi_t(x_t)$  is decreasing in all costs, i.e., decreasing in  $c$ ,  $h$ , and  $r$ .*

The cost parameters are the only parameters in the inventory-price control model for which the value function satisfies monotonicity. An increase in the market potential  $\alpha$  leads to an increase in demand, which can increase revenue but can also increase potential shortage costs – a nonmonotonic effect. An increase in the elasticity of demand  $\beta$  also produces a nonmonotonic effect: it can decrease demand, for a given price  $p$ , thereby decreasing revenue *but* can also decrease the potential shortage costs. The nonmonotonic behavior of the payoff function in these other parameters will limit our ability to provide neat comparative statics results.

The single-period value function can be shown to exhibit increasing differences in the shortage cost  $r$  and the negative holding cost,  $-h$ . From Milgrom and Shannon (1994), a twice

continuously differentiable function  $f(x, t)$  increasing differences in  $(x, t)$  is equivalent to  $\frac{\partial^2 f}{\partial x_i \partial t_j} \geq 0$ , for  $i = 1, \dots, n, j = 1, \dots, m$ . We have

$$\frac{\partial}{\partial r} L_p = -\beta + \beta P(\tilde{d} \leq \tilde{s}) \leq 0$$

$$\frac{\partial}{\partial r} L_y = -\lambda + \lambda P(\tilde{d} \leq \tilde{s}) \leq 0.$$

The loss function,  $-L(p, y)$  exhibits increasing differences in  $r$ . Also,

$$\frac{\partial}{\partial h} L_p = \beta P(\tilde{d} \leq \tilde{s}) \geq$$

$$\frac{\partial}{\partial h} L_y = \lambda P(\tilde{d} \leq \tilde{s}) \geq 0,$$

and we see that the loss function,  $-L(p, y)$  exhibits decreasing differences in  $h$  or *increasing* differences in  $-h$ . These are intuitive results. Increasing differences, or complementarity, implies that for a higher value of a given parameter it is in some sense “better” to also have more of  $p$  and more of  $y$ . As the shortage cost  $r$  increases, it is preferable to have a higher price  $p$  (to dampen demand and reduce losses) and a higher order level  $y$  (to augment supplies and hence offset the now more-expensive shortages). Conversely, for a higher holding cost  $h$ , it is better to have a lower price  $p$  (to encourage demand and reduce possible left-over inventory, which is now more expensive to store) and a lower order level  $y$  (to, again, reduce the amount of left-over inventory at the end of the period).

Similar checks reveal that the single-period value function  $g(p, y)$  exhibits increasing differences in the market potential  $\alpha$  and decreasing differences in the elasticity of demand  $\beta$ . Again, this is in keeping with our economic intuition: in the case of stronger demand ( $\alpha \uparrow$ ), it is better to order more, i.e., have a higher order level  $y$ , and at least not to decrease price  $p$  (possibly raising it to offset potential shortage losses). Facing a lower demand elasticity, price  $p$  is increased (both to extract additional revenue in the face of less price-sensitive demand and to avoid additional shortage losses); the order-level  $y$  is also increased (to meet higher demand, thereby garnering additional revenues, and also to avoid higher shortage costs).

In conjunction with the results above, supermodularity of the per-period value function in  $(p, y)$  would suffice for monotone comparative statics to obtain. However, as established in Propositions 2.2 and 2.7, our per-period payoff function is *submodular* in  $(p, y)$ . While it is desirable, for instance, to increase both  $p$  and  $y$  in the face of increased shortage costs  $r$ , the overall impact of an increase in both can be deleterious on the intermediary’s payoff function. The submodularity in  $(p, y)$  of the value function is a defining feature of the

joint inventory and pricing control model, and it necessarily complicates comparative statics predictions. Whether it is better to increase price  $p$  or to increase the order level  $y$  when facing a higher shortage cost  $r$  cannot be discerned from relationships discussed above, since they are contradictory: increasing differences of  $g(p, y, r)$  suggests it is better to have more of both  $p$  and  $y$ , whereas the submodularity result says otherwise, warning that a simultaneous increase in both will have a negative impact. These trade-offs are explored in the more detail in the numerical study in Section 2.4.

Before concluding this section, we provide an additional positive result, in the context of uncertainty in the model. Specifically, we are interested in how shifts in our supply uncertainty, as characterized by shifts in the mean or variance, impact the optimal control policy. The proofs in below draw on results presented in Athey (2002) and Smith and McCardle (2002).

The demand and supply shocks in our model are represented by the bivariate normal density. The univariate normal density satisfies the monotone likelihood ratio property (MLR) with respect to its mean but *not* with respect to its variance (Karlin and Rubin (1956)). To see why this is so, consider the implication of satisfying the MLR. For a density to satisfy the MLR with respect to one of its parameter, it must be such that an increase in the parameter makes a good outcome, i.e., a higher realization of the random variable, more likely without making a bad outcome more likely. A (positive) shift in the mean makes a higher outcome more likely without increasing the chances of a bad outcome. However, a (positive) shift in the variance makes both good (higher) and bad (lower) outcomes more likely. The bivariate normal density also satisfies MLR with respect to either variate's mean (or both), but only for  $\rho \geq 0$ . In Proposition 2.10 below we present a result specific to our model and a more general version. The general result holds for distributions satisfying the MLR in a given parameter. For instance, the exponential family of distributions satisfies MLR. The exponential family includes the normal (with fixed variance), binomial, Poisson, and Gamma (Karlin and Rubin 1956).

First, we introduce three assumptions needed in the proof of Proposition 2.10. Assumption A4 is satisfied if assumptions A2 and A3 are both met in our model. The assumptions are essentially restrictions on the primitive value function and the joint density of supply and demand to ensure that each is log-supermodular, where the primitive value function in our model is given as

$$v(p, y, \epsilon, \xi) = p_t \tilde{q}_t - c(y_t - x_t) - \max \left\{ h \left( \tilde{s}_t(y_t) - \tilde{d}_t(p_t) \right), r \left( \tilde{d}_t(p_t) - \tilde{s}_t(y_t) \right) \right\} \\ + \omega \left( (\hat{p}_t - p_t) \tilde{q}(p_t, y_t) + (\tilde{q}(p_t, y_t))^2 / (2\beta) \right)$$

A2. (“Strong Enough Demand”): For any  $\epsilon$ , we have  $\tilde{d}(p) - \beta(p + h) \geq 0$ . Note that

this condition is satisfied for restrictions on the distribution of  $\epsilon$  that guarantee  $\alpha \geq \beta(2p + h)$ . Without restrictions on the values  $\epsilon$  can take one - and, hence, with the possibility of demand being zero - the condition is more succinctly stated as simply  $\beta = 0$ , i.e., inelastic demand.

- A3. (*“High Enough Value”*): Both  $v \geq (p+h+\epsilon)(\tilde{d}(p)-\beta(p+h))$  and  $v \geq (p+r)(\tilde{s}(y)+\beta r)$ , where  $v$  is the primitive per-period payoff function,  $v(p, y, \epsilon, \xi)$ .
- A4. (*Log-Supermodularity of the Per-Period Payoff Function*): The per-period payoff function  $g(p, y, \epsilon, \xi)$  is log-supermodular; i.e., the primitive function  $v(p, y, \epsilon, \xi)$  is log-supermodular and the joint density  $\varphi(\epsilon, \xi; \theta)$  is also log-supermodular, i.e.,  $\varphi$  satisfies the monotone likelihood ratio order (MLR) in  $\theta$ .

In our model, assumptions A2 and A3 ensure log-supermodularity of the primitive value function  $v(p, y, \epsilon, \xi)$ , a property that allows us to establish that the optimal solution,  $\arg \max_{(p_t, y_t) \in \mathbb{R} \times [x_t, x_t+z]} \Pi_t(x_t)$ , is increasing in the mean of the demand shock  $\mu_\epsilon$ . (See Step 1 of the proof of Proposition 2.10.) The log-supermodularity property is akin to assuming that the function exhibits increasing relative returns.

From Assumption A2 we see that the special case of completely inelastic demand ( $\beta = 0$ ) is of interest. Assumption A4 summarizes the general conditions under which monotone comparative statics will hold with relation to parameters of the density function our model uncertainty (see e.g., Athey 2002).

**Proposition 2.10** (i) For “strong enough demand” and “high enough value,”  $\arg \max_{(p_t, y_t) \in \mathbb{R} \times [x_t, x_t+z]} \Pi_t(x_t)$  is increasing in the mean of the demand shock  $\mu_\epsilon$ , for  $\omega = 0$ , i.e., under profit maximization. (ii) More generally, under assumption 3 above,  $\arg \max_{(p_t, y_t) \in \mathbb{R} \times [x_t, x_t+z]} \Pi_t(x_t)$  is increasing in the parameter  $\theta$  of the joint density of supply and demand uncertainty,  $\varphi(\epsilon, \xi; \theta)$ , for  $\omega = 0$ .

The above proposition does not hold for the case of welfare maximization ( $\omega = 0$ ). (The partial derivative of the log of the primitive function,  $\log v(p, y, \epsilon)$  is *always* negative, and, hence, our primitive value function  $v(p, y)$  does not satisfy log-supermodularity.) Under welfare maximization, the impact of increasing the price  $p$  and the order level  $y$  is amplified - an increase in price is an even worse affect due to the additional consumer surplus term (which is negatively impacted), and the increase in  $y$  still increases both the order cost and the potential storage cost.

## 2.4 Application to the Metropolitan Water District

In this section we discuss the application of the joint inventory-price control model to the largest water intermediary in California, the Metropolitan Water District of Southern California (MWD). We conduct a detailed numerical study, examining the impact of differing costs, shifts in uncertainty, and changes in the correlation between supply and demand. Although it seems natural that MWD as a public intermediary would be interested in welfare-maximization, versus profit maximization, we first explore the intermediary's behavior under profit maximization. This allows us to compare the price and order-level behavior to previous modelling efforts, which have assumed profit-maximization in view of their generally application in a retail setting. We then compare social welfare maximizing behavior to profit maximizing behavior. As intuition would suggest, optimal prices under welfare maximization are lower and optimal order-levels are higher, in comparison to those under profit maximization. We begin with a discussion of MWD's current operations, on which we base our parameterization of the model.

### 2.4.1 Background on MWD

MWD contracts with the State Water Project for close to 2 maf of water a year. Although MWD holds other supply contracts, the SWP contract is the largest and currently accounts for over two-thirds of MWD's contractual supply. In dry years or when reservoir levels are low, the SWP deliveries are cut by a percentage, as determined by the state's Department of Water Resources (DWR). Figure 2.2 (b) shows MWD's total estimated supply over a 20-year period, from 1986 to 2005.<sup>78</sup>

Demand for water in Southern California is higher during dry years; our regression analysis indicated that the uptake in demand is present not just for years categorized as "critical" or "dry," but also for those categorized as "below normal" in the DWR's five-point year-type index. (The other two year classifications in the index are "above normal" and "wet.") We would therefore expect the correlation coefficient for supply and demand to be negative. However, the DWR does not simply divert water from the major river systems in northern

---

<sup>7</sup>We have assumed that the other contractual suppliers have instituted the same cuts in supply as DWR; that is, if DWR cuts SWP supply by 25%, then it is assumed that the remaining one-third of MWD's contractual supply from other sources was also cut by 25%.

<sup>8</sup>Under this assumption, the supply variability is overestimated. In actuality, MWD's additional contractual supply is not perfectly correlated with the SWP supply and, hence, at times where SWP imposes cuts in deliveries to MWD, the remaining contracts may still be filled.

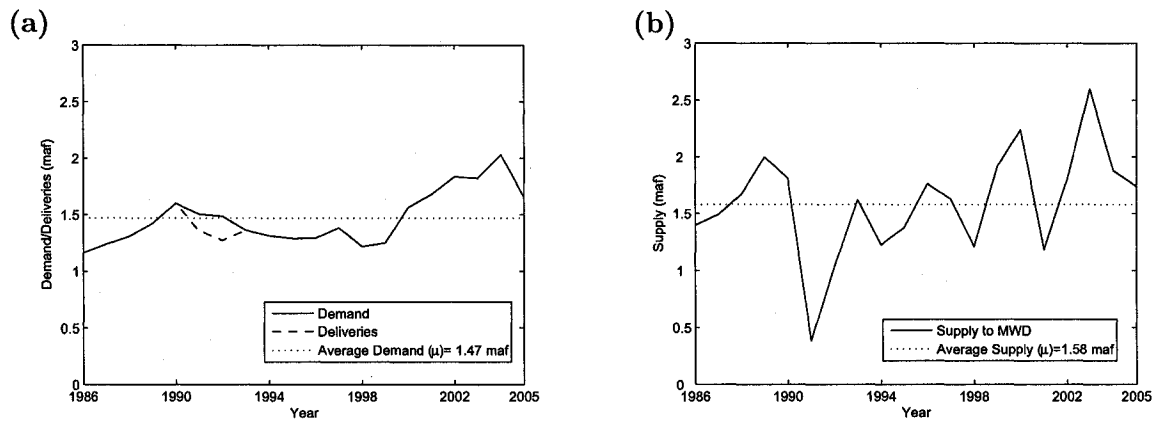


Figure 2.2: 20-Year Demand and Supply

California to provide to its contractors, such as SWP. Rather, it operates a system of reservoirs, which it uses to offset supply shortfalls in dry years. Recognizing that demand by all sectors will increase in a dry year, the SWP will not necessarily cut supplies. If the reservoir levels are high in a given dry year (which may be the case if, for example, the previous year was wet), the SWP may opt to augment river supply by releasing water from its reservoirs. Faced with low reservoir levels, as may be the case during a multi-year dry spell, DWR may not have the flexibility to release additional water from the reservoirs, resulting in supply cuts. As a consequence, Figure 2.3 (a) shows that the correlation coefficient between MWD demand and supply (from SWP) is positive. Figure 2.3 (b) shows that, in contrast, the “true” correlation between supply (from the major river systems) and demand is indeed negative.

The early 1990s was the culmination of a six-year drought in California, the worst drought on record in a 100-year period. In 1991 and 1992, supply levels fell drastically, and MWD implemented rationing. Rationing was instituted in accordance with MWD’s five-stage Incremental Interruption and Conservation Plan (IICP), where Stage 3 required that all member agencies reduce their stated demands to 10% below 1990 levels; Stage 4 required a 20% reduction from 1990 levels, and Stage 5, declared in April 1991, 30%.<sup>9</sup> (See Appendix B for further details of the IICP.) The line “Demand” in Figure 2.2 (a) estimates actual demand

<sup>9</sup>MWD, letters to the Board containing the subject, “Incremental Interruption and Conservation Plan,” Nov. 13, 1990 to Sept. 9, 1993, [www.mwdh2o.com](http://www.mwdh2o.com), retrieved from [<http://edmsidm.mwdh2o.com/idmweb/search.asp>] on Nov. 10, 2006, with “Subject contains:” set equal to “Incremental Interruption and Conservation Plan.”

vs. deliveries, during the period when rationing was in place (from 1991 to 1993), where all demands by individual member agencies have been adjusted back to the 1990 levels whenever the constraint that demand be less than 10%, 20% or 30% of 1990 demand levels, depending on the level of rationing in place for that month, are binding for the agency.

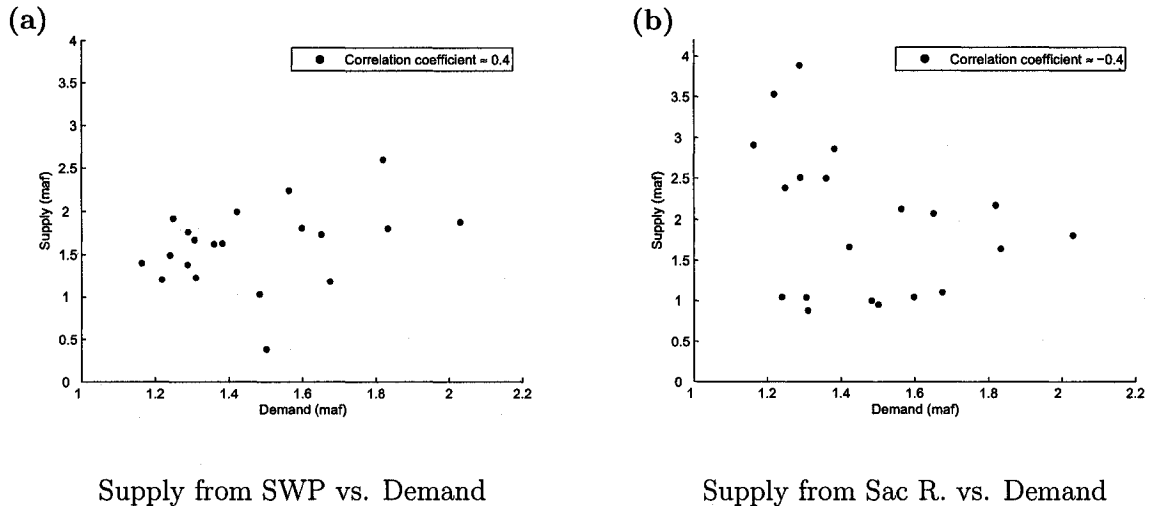


Figure 2.3: Supply-Demand Correlation

The estimates for  $\alpha$  and  $\beta$ , the market potential and the elasticity-of-demand parameter, respectively, in our linear demand model were estimated using a multiple linear regression, with  $R^2 = 0.8652$ . The model accounted for seasonality and population growth in Southern California, using the population growth in San Diego county, as reported in the U.S. census, as the proxy for regional population growth. Not all areas in the MWD service area are growing as rapidly as San Diego county; however, there is a strong regional growth trend. (San Diego is currently the largest user of MWD water, accounting for roughly 30% of their total deliveries in 2005.) The seasonality factors are accounted for by the dummy variables  $d_{crit}$ , which is equal to one if the year-type is “dry,” “critical,” or “above normal,” and zero otherwise, and  $d_{wet}$ , which is equal to one if the year type is “wet,” and zero otherwise. The variable  $cumpopgrowth$  tracks the cumulative population growth each year in San Diego County.<sup>10</sup> The demand model is given as  $d(p) = \alpha - \beta p + d_{crit} \gamma_c + d_{wet} \gamma_w + cumpopgrowth g$ , where  $\alpha$  is the market potential,  $\beta$  is the price elasticity, and  $\gamma_c$ ,  $\gamma_w$ , and  $g$  are weights on the dummy variables and the population variable, respectively. Estimates for the parameters for the model with standard errors (in parentheses) reported for the 5% significance are reported in

<sup>10</sup>U.S. CENSUS BUREAU (1986–2005) “Annual Estimates of the Population for Counties,” [www.census.gov](http://www.census.gov).

Table 2.1.

For the simulation runs, population growth was not included in the model, and, hence the dummy variable *cumpopgrowth* below was assumed equal to zero in all scenarios; similarly, the dummy variables *dcrit* and *dwet* were set to zero. Under these assumptions, the demand model reduces to  $d = \alpha' - \beta p + \epsilon$ , as in our derivations, where  $\alpha'$  is the model-estimated  $\alpha$  plus the most recent population factor, which is held constant in our simulations.

$\alpha$	$\beta$	$\gamma_c$	$\gamma_w$	$g$
2,939,200 af (160,890)	2,600 af/\$ (740)	17,900 (57,620)	-122,100 (66,350)	4.4 (0.77)

Table 2.1: Demand Estimation Parameters

The unit cost of supplying water,  $c$ , is estimated using MWD's records for water supply in 2005, as their total annual costs, minus the fixed costs (as charged through their Capacity Charge and Readiness-to-Serve), divided by the total water delivered.<sup>11</sup> The holding cost for water, which is negligible for an additional af of water in a reservoir but is estimated to be \$140/af for groundwater storage, is the weighted average of these two, where the weights are the percentage of existing reservoir capacity (64%) and groundwater capacity (46%), respectively.

Description	Symbol	Value
Market potential	$\alpha$	2.9392 maf
Elasticity of demand	$\beta$	2.6 taf
Unit order cost	$c$	\$400
Holding (storage) cost	$h$	\$0 to \$64/af
Shortage cost	$r$	\$0 to \$5,885/af
Capacity	$\bar{x}$	3.26 maf
Loss coefficient	$\lambda$	1
Discount factor	$\delta$	0.95
Joint dist. of supply and demand	$\varphi(\epsilon, \xi)$	Bivariate normal
Std. dev. of demand	$\sigma_\epsilon$	2.372 taf
Std. dev. of supply	$\sigma_\xi$	4.632 to 8.895 taf
Correlation coefficient	$\rho$	-0.4 to 0.4

Table 2.2: Parameter Values for MWD

<sup>11</sup>MWD, letter to the Board, 10 Jan. 2006, Board Letter Archives, [www.mwdh2o.com](http://www.mwdh2o.com), retrieved from [<http://edmsidm.mwdh2o.com/idmweb/cache/MWD%20EDMS/003678343-1.pdf>] on Nov. 10, 2006.



The shortage costs come from a study commissioned by the San Diego County Water Authority, an MWD member, and conducted by the economic consulting firm CIC Research, Inc. The study examines the consequences of incremental water supply cuts of 20%, 60%, and 80% over a two or six-month period, as might occur during a supply interruption caused by an earthquake, or, possibly, a severe drought.<sup>12</sup> The study assessed the costs associated with a two-month, or six-month supply disruption, at varying levels of cutbacks in deliveries (20%, 60%, and 80%). The estimated shortage cost, per af, is \$5,885, more than ten times the estimated unit cost of water. As MWD does not currently use an estimate of shortage costs to inform their rate setting, we varied shortage costs in our simulation, using values of \$0, \$5,885, and \$2,500. The total capacity is as reported by MWD in their 2005 Regional Urban Water Management Plan. Table 2.2 above reports these parameter values.

As discussed earlier, the time intervals,  $t$ , in the model should be selected in accord with the actual schedule for decisions by the intermediary, whether it be daily, weekly, quarterly, or annually. MWD sets rates annually, and, hence, we have chosen an annual time step:  $t = 1$  corresponds to year one.<sup>13</sup> The question of what time horizon to use is a somewhat trickier issue. We adopt a 10-year horizon, finding, as have Federgruen and Heching (1999), that the year one policy, with this horizon, is stable.

## 2.4.2 Shortage and Holding Cost Variability

Storage, or holding, costs for water are likely to change over time with an increase in (more costly) subsurface storage, the changing costs of technologies associated with subsurface storage, and the cost electricity (used to pump water out of subsurface storage).<sup>14</sup> Shortage costs are also subject to change over time, since the value of water (and hence the cost of foregoing a unit of it) is tied to variable economic factors such as productivity and industry types. Also, as we noted earlier, there is no consensus on the existing shortage cost. Instead, we specify what we consider a representative range of shortage costs. Below we explore the impact of changes in the holding and shortage costs on the optimal policies.

In Figure 2.4, holding shortage costs constant at their mid-value of \$2,500/af, we see that as

---

<sup>12</sup>SAN DIEGO COUNTY WATER AUTHORITY (Dec. 28, 1993, Revised Dec. 28, 1994) "The Economic Impact on San Diego County of Three Levels of Water Delivery: 80, 60, or 40 Percent Occurring for Two Months or Six Months," By CIC Research, Inc.

<sup>13</sup>Decisions about contracts, or order levels, are made on an ongoing basis at MWD, so the reduction of the decision to a single annual review represents a simplification in this respect.

<sup>14</sup>Another foreseeable change in storage cost might be a "security surcharge," i.e., the cost of protecting water supplies from contamination.

holding costs increase, the optimal order-up-to-level,  $y^*$ , decreases. The optimal list price  $p^*$  remains constant, even for a 10 fold increase in the holding cost; however, we see that, for  $x > y^*$ , the price discounts are greater in magnitude and increase more rapidly the higher the holding cost. We see similar behavior, in Figure 2.5, where  $r = \$0$ ; the difference in this case is that with no shortage costs, the optimal order-up-to-levels are lower and the price discounting is even more dramatic.

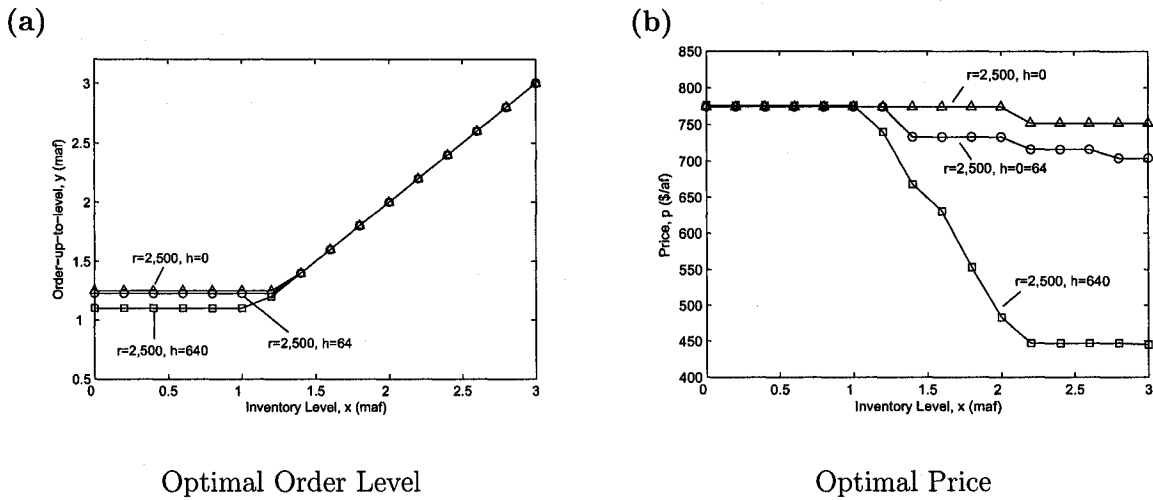


Figure 2.4: Sensitivity of Policy to Holding Costs ( $r = \$2,500$ )

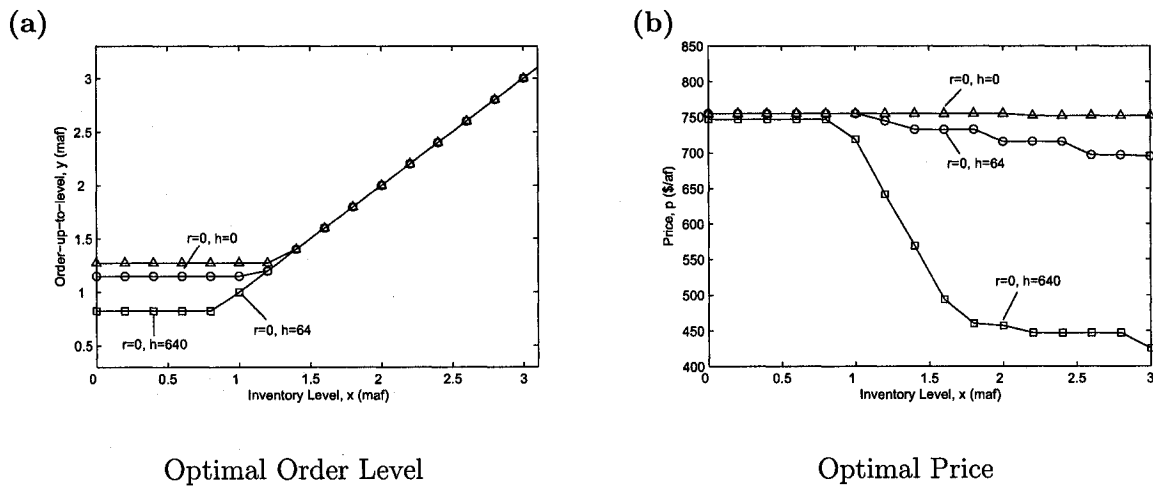


Figure 2.5: Sensitivity of Policy to Holding Costs ( $r = \$0$ )

For varying shortage costs, with holding cost  $h$  set equal to  $\$0$ , we observe that the optimal order-up-to-levels remain almost the same; however, prices increase with shortage costs – up

to a threshold at least. When the shortage cost jumps from \$2,500/af to \$5,885/af, the price does not increase any further. Overall, we observe that the policy adjustments associated with increases in holding versus shortage costs manifest themselves quite differently. The optimal response to a change in the holding costs dictates a change in the optimal order-up-to-level  $y^*$ , whereas changes in the shortage costs dictate a change in the optimal price  $p^*$ . (The holding cost changes do impact the price as well, although only in cases where  $x > y^*$  and price discounts are offered below the list price.)

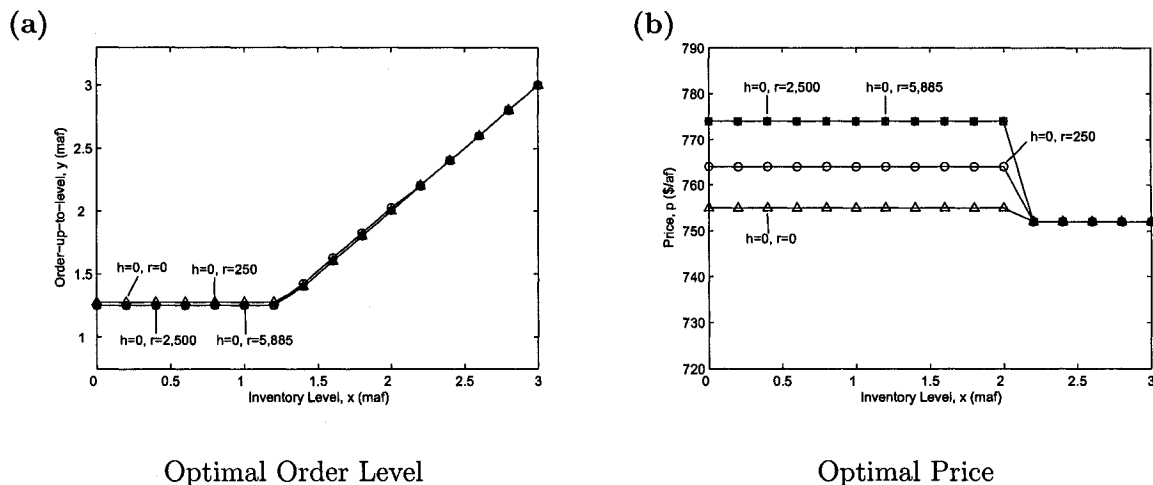


Figure 2.6: Sensitivity of Policy to Shortage Cost

### 2.4.3 Shifts in Uncertainty

The supply uncertainty that MWD faces depends on both the weather and the mix of supply contracts it holds, i.e., due to the prioritization of water rights some supplies are considered “firm” whereas others are met only in years of above average streamflows. The impacts of climate change on both the volume and variability of annual precipitation and on the temperature (which in turn impacts the timing of the snowpack melt and, hence, the timing of available supplies) are a chief concern of water managers. On the demand front, population growth and economic growth have both the potential to shift the mean and the variance of the demand uncertainty.

In Section 2.3, we established comparative statics results given several assumptions. These included the requirement that the demand elasticity  $\beta$  be zero, a condition which is not met here. Nonetheless, we do find in our simulations that the optimal order level  $y^*$  and the optimal price  $p^*$  are increasing in the mean of the demand uncertainty  $\mu_\epsilon$ . (See Figure 2.7

below, where the optimal order-up-to-level  $y^*$  and the optimal price  $p^*$  are increasing in the mean of the demand uncertainty  $\mu_\epsilon$ . The supply uncertainty is negligible:  $\sigma_\xi = 0.01$ ;  $h = \$0/\text{af}$  and  $r = \$2,500/\text{af}$ .)

In Figures 2.8 and 2.9, we investigate the impact on the optimal policy of an increase in the variance of demand and supply, respectively. The optimal order-up-to-level,  $y^*$ , and the optimal list price,  $p^*$  are increasing in the demand uncertainty,  $\sigma_\epsilon$ . (The supply uncertainty in Figure 2.8 is negligible;  $\rho = 0.4$ ,  $h = \$0/\text{af}$ , and  $r = \$2,500/\text{af}$ .) For pure supply uncertainty, both  $y^*$  and  $p^*$  are increasing in the uncertainty  $\sigma_\xi$ . (In Figure 2.9,  $\rho = 0.4$ ,  $h = \$0/\text{af}$ , and  $r = \$2,500/\text{af}$ .)

Figure 2.10 shows results for a joint increase in supply and demand uncertainty. As we noted in Section 2.3, comparative statics results for the variances of the supply and demand uncertainty do not hold for the normal density since it fails to satisfy the monotone likelihood ratio order in  $\sigma_\epsilon$  or  $\sigma_\xi$ . We see in Figure 2.10 that although it appears that the optimal order level  $y^*$  is increasing in the joint supply and demand uncertainty that, in fact, as they become extremely high the optimal order level  $y^*$  actually drops to zero. In Figure 2.10,  $\rho = 0.4$ ,  $h = \$0/\text{af}$ , and  $r = \$2,500/\text{af}$ .

Figure 2.11 shows the impact of the increases in demand and supply uncertainty on the expected payoff function. The payoff function is increasing in the demand uncertainty, for  $h = 0$ ; however, for  $h = 640$ , payoff levels are *decreasing* in the demand uncertainty. In the case of pure supply uncertainty (and  $h = 0$ ), the payoff function is uni-modal. In (b), with two-sided uncertainty and *positively* correlated supply and demand ( $\rho = 0.4$ ), the payoff function is increasing in the joint uncertainty. In the case of *negatively* correlated supply and demand uncertainty, the payoff function is bi-modal. Finally, Figures 2.12 and Figure 2.13 examine the effects of negative versus positive correlation between supply and demand. The optimal order-up-to-level  $y^*$  is decreasing in  $\rho$ . The optimal list price  $p^*$  does not shift significantly. The timing of the price discount, in the case  $y > x^*$ , is increasing, i.e., earlier, in  $\rho$ . The payoff function is increasing in  $\rho$ .

#### 2.4.4 Policy Implications

Welfare maximization results in a higher order-up-to-level,  $y^*$ , and a lower price,  $p^*$ , as established in Section 2.3. The price shift is significant: under welfare maximization, the price drops by  $\$300/\text{af}$ , or nearly half, to be closer to  $\$400/\text{af}$ . MWD's Figure 2.14 below shows the optimal order level and price under both profit maximization and welfare maximization. The current MWD price ( $\$/\text{af}$ ) of  $\$478$  is above our optimal welfare maximization price. The

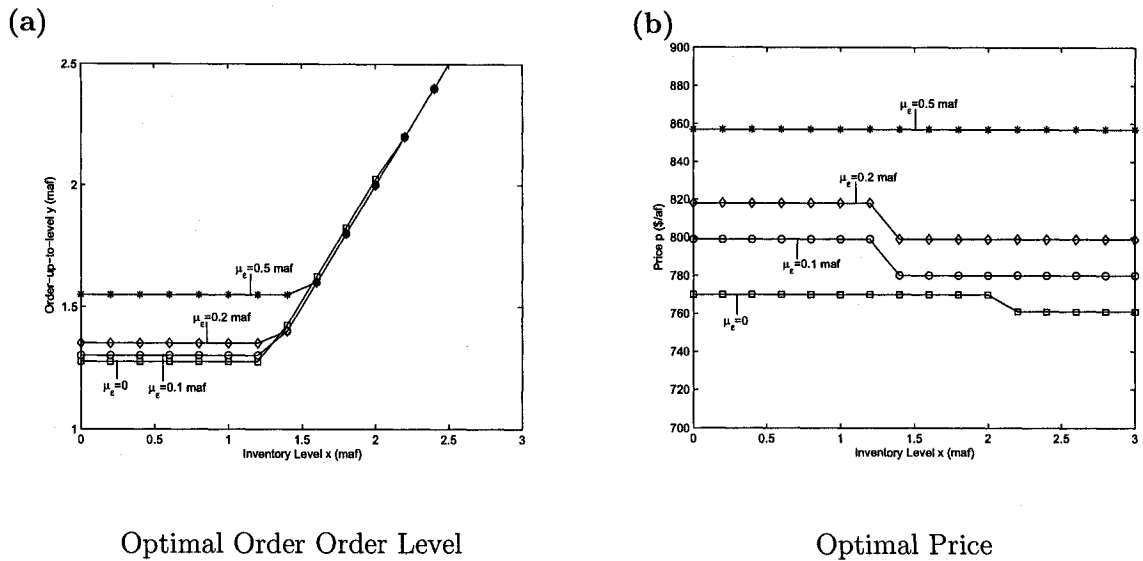


Figure 2.7: Sensitivity of Policy to Shifts in the Mean of Demand

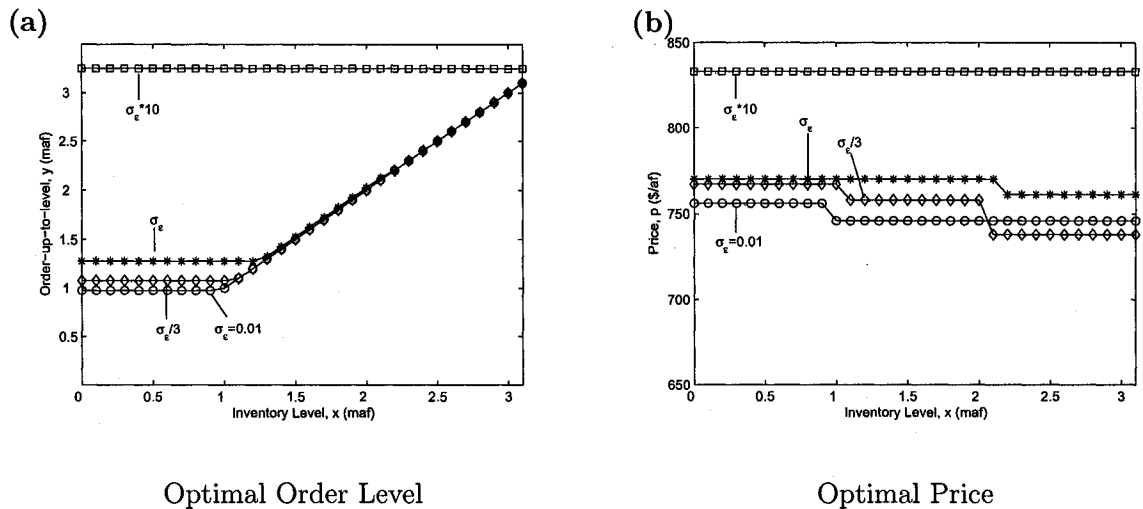


Figure 2.8: Sensitivity of Policy to Pure Demand Uncertainty

difference may be due in part to MWD's assessment of future capital needs (and inclusion of part of this cost in the per-unit charge for water); it may also suggest sub-optimal, i.e., too low, order, or contractual, levels. The current MWD order level, or contractual supply level, is harder to assess given interannual changes in the options contracts that they hold.

As discussed earlier, the "true" correlation between supply and demand, i.e., between annual streamflow and urban water demand, is negative. However, the correlation between SWP

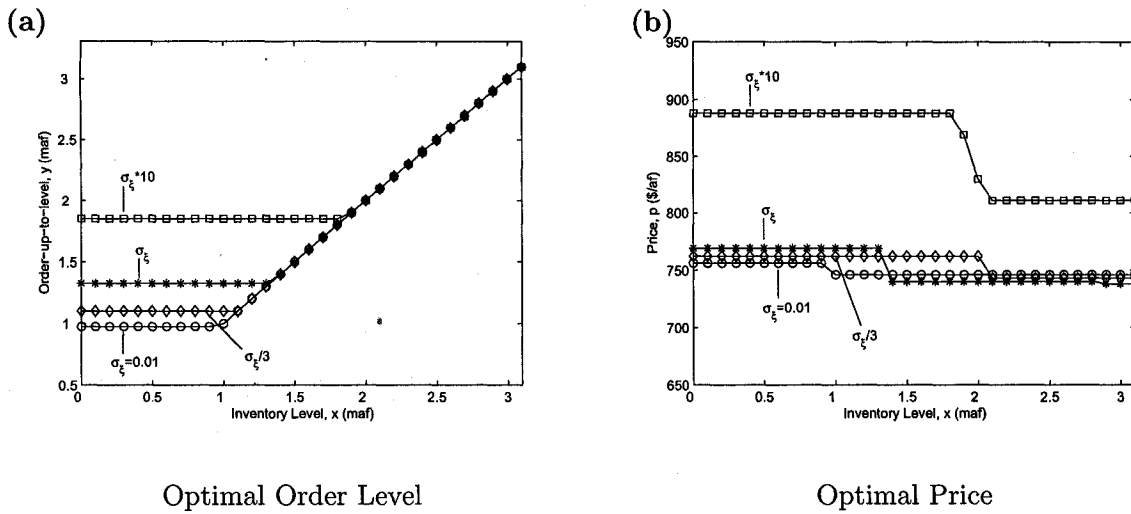


Figure 2.9: Sensitivity of Policy to Pure Supply Uncertainty

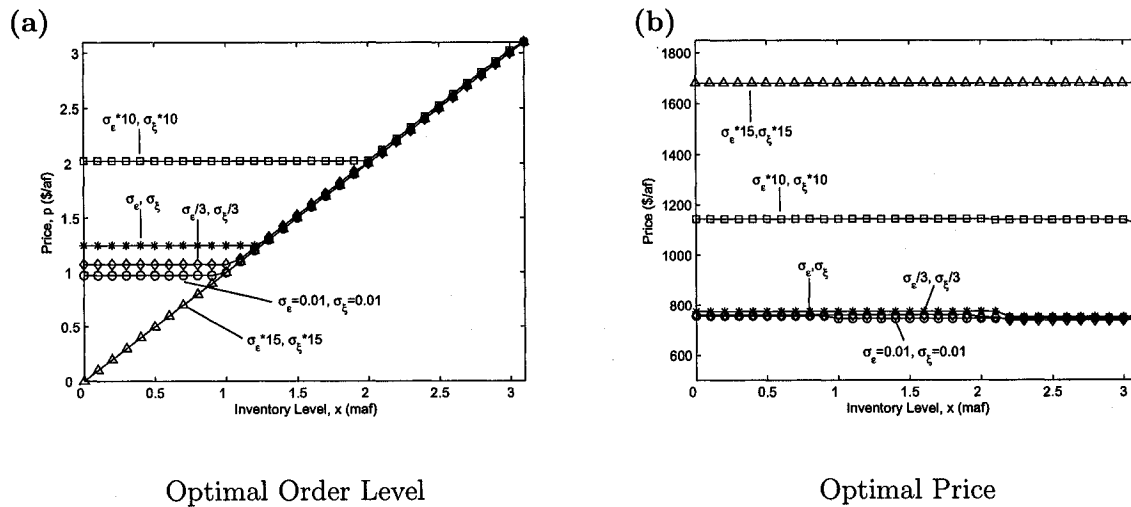


Figure 2.10: Sensitivity of Policy to Joint Uncertainty

supply to MWD and MWD's demand is actually positive. In Figure 2.15 below we see that, under welfare maximization, the optimal price and order level are decreasing in the correlation coefficient. MWD would be able to more effectively manage its operations, i.e., set price and order levels, if it could manage its uncertainty directly rather than being subject to SWP policy regarding when to store and when to release reservoir water. One way to achieve this would be through a contract specifying that MWD controls a percentage of SWP storage (where the percentage could be based on the percentage of total SWP supplies that

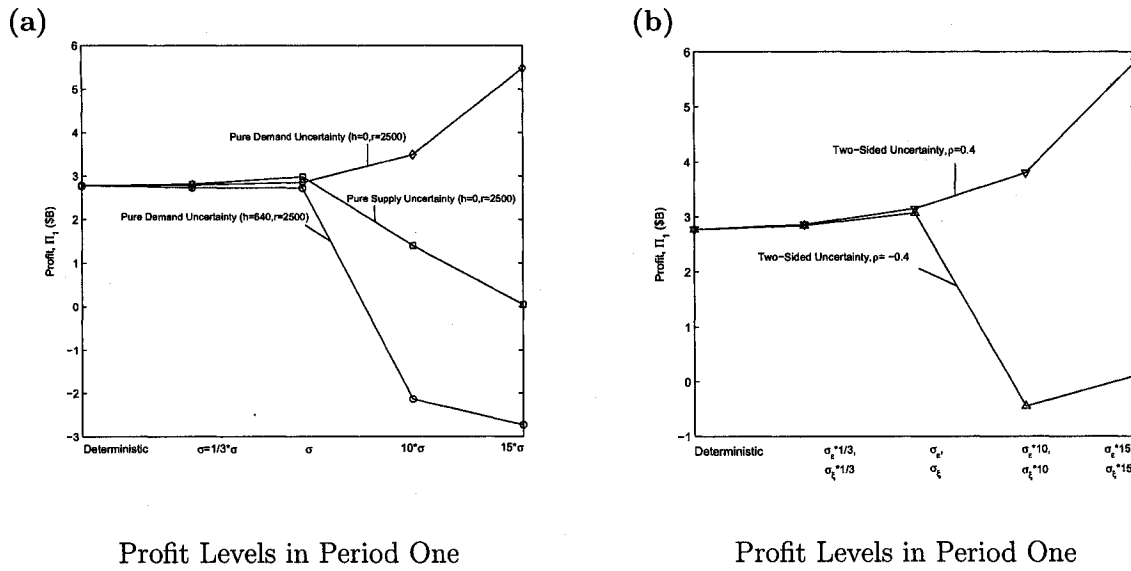
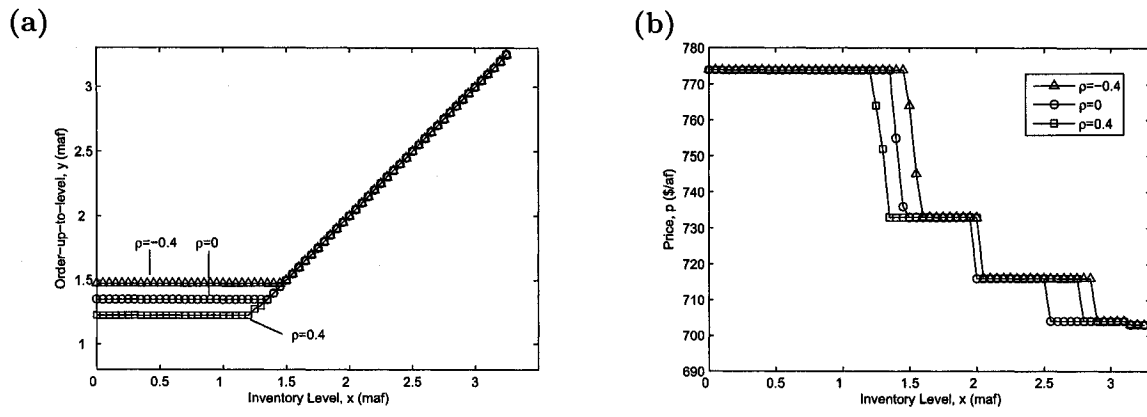


Figure 2.11: Sensitivity of Payoff to Uncertainty

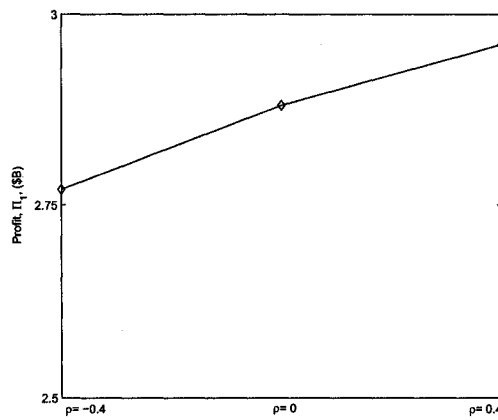
MWD currently contracts for, approximately 50%). This would allow MWD to optimize its operations in accord with the true supply instead of the possible sub-optimal SWP supply policies.



Optimal Order Level

Optimal Price

Figure 2.12: Sensitivity of Policy to the Correlation Coefficient  $\rho$



Profit Levels in Period One

Figure 2.13: Sensitivity of Payoff to the Correlation Coefficient  $\rho$

## 2.5 Conclusion

We have extended the joint inventory and pricing control model to a setting with two-sided uncertainty, no backlogging, and limits on both the supplier's capacity and the intermediary's order level. This setting is realistic for an intermediary who faces the following conditions: (1) both upstream and downstream uncertainty, (2) demand that cannot be carried over to future periods, and (3) fixed short-run capacity and supply-level constraints. In addition, we



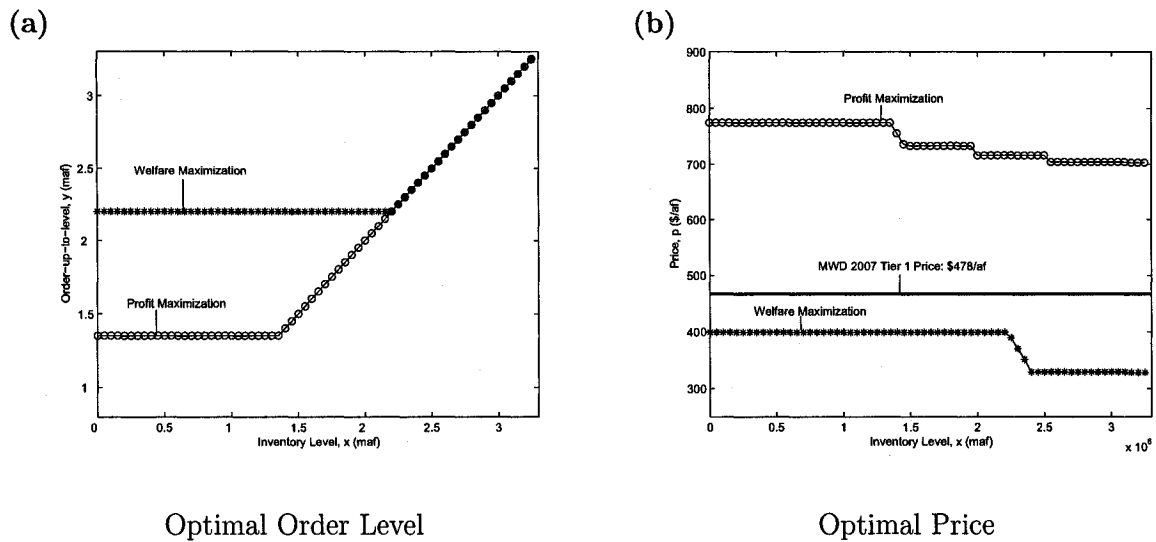
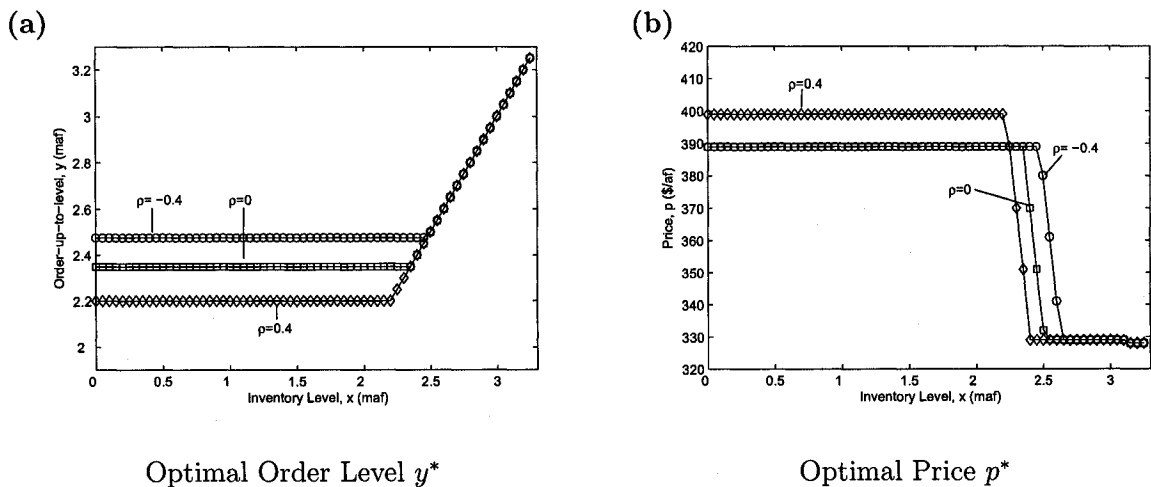


Figure 2.14: Sensitivity of Policy to Profit vs. Welfare Maximization

Figure 2.15: Sensitivity of Policy to the Correlation Coefficient  $\rho$  (Welfare Maximization)

have expanded the traditional application of inventory and pricing control models to consider allocative efficiency in terms of a welfare-maximizing policy. We consider results for both profit-maximizing and welfare-maximizing intermediaries and demonstrate the intuitively appealing result that the optimal order level  $y^*$  is higher under welfare maximization, while the optimal price  $p^*$  is generally lower.

The application of inventory and pricing control models to public pricing problems has not, to date, received significant attention. The advantages of such applications over existing

public pricing models include the ability to capture storage dynamics and to model ordering decisions with fixed short-run capacity. As noted in the introduction, the extant public pricing literature provides models that treat joint pricing and *capacity* decisions, not pricing and *ordering* decisions, thereby ignoring storage. Our public pricing model for water may also be applicable to other natural resources, such as oil and natural gas, and even to electricity, where the storability of electricity – via dams, fly wheels, and fuel cells – call for a framework other than the traditional peak-load pricing models.

Finally, the application of our model to the large water intermediary, the Metropolitan Water District of Southern California, illustrates the potential of such models to inform both operations and policy regarding such institutions.

## Appendix A: Proofs

**Proof of Proposition 2.1.** By definition, the expected demand served is

$$Q(p, y) = \int_{\mathbb{R}} \left( \int_{-\infty}^{\min\{\lambda y + \xi, \bar{x}\} - (\alpha - \beta p)} (\alpha - \beta p + \varepsilon) \varphi(\varepsilon, \xi) d\varepsilon + \int_{\min\{\lambda y + \xi, \bar{x}\} - (\alpha - \beta p)}^{\infty} \min\{\lambda y + \xi, \bar{x}\} \varphi(\varepsilon, \xi) d\varepsilon \right) d\xi.$$

Hence, by differentiation (using Leibniz rule) we obtain that  $Q$  is increasing in price  $p$ ,

$$Q_p(p, y) = -\beta P(\tilde{d}_t(p) \leq \tilde{s}_t(y)) < 0,$$

and that  $Q$  is increasing in the order-up-to level  $y$ ,

$$Q_y(p, y) = \lambda P(\tilde{s}_t(y) < \min\{\tilde{d}_t(p), \bar{x}\}) > 0,$$

so that the cross-derivative of  $Q$  with respect to  $(p, y)$  becomes

$$Q_{py}(p, y) = -\beta \lambda \int_{-\infty}^{\bar{x} - \lambda y} \varphi(\lambda y + \xi - (\alpha - \beta p), \xi) d\xi < 0.$$

The last inequality implies the submodularity of  $Q$  with respect to  $(p, y)$ . We now consider the consumer surplus  $\text{CS}(p, y)$ , which can be written in the form

$$\text{CS}(p, y) = E \left[ \left( \frac{\alpha}{\beta} - p \right) \tilde{q}(p, y) + \frac{\tilde{\varepsilon} \tilde{q}(p, y)}{\beta} - \frac{\tilde{q}^2(p, y)}{2\beta} \right].$$

Hence, the derivative of expected consumer surplus with respect to  $p$  becomes

$$\text{CS}_p(p, y) = \frac{\partial}{\partial p} E \left[ \left( \frac{\alpha}{\beta} - p \right) \tilde{q}(p, y) \right] + \frac{\partial}{\partial p} E \left[ \frac{\tilde{\varepsilon} \tilde{q}(p, y)}{\beta} \right] - \frac{\partial}{\partial p} E \left[ \frac{\tilde{q}^2(p, y)}{2\beta} \right].$$

The second term can be computed as follows:

$$\begin{aligned} \frac{\partial}{\partial p} E \left[ \frac{\tilde{\varepsilon} \tilde{q}(p, y)}{\beta} \right] &= \frac{1}{\beta} \frac{\partial}{\partial p} \left( \int_{\mathbb{R}} \left[ \int_{-\infty}^{\min\{\lambda y + \xi, \bar{x}\} - (\alpha - \beta p)} \varepsilon (\alpha - \beta p + \varepsilon) \varphi(\varepsilon, \xi) d\varepsilon \right. \right. \\ &\quad \left. \left. + \int_{\min\{\lambda y + \xi, \bar{x}\} - (\alpha - \beta p)}^{\infty} \varepsilon \min\{\lambda y + \xi, \bar{x}\} \varphi(\varepsilon, \xi) d\varepsilon \right] d\xi \right) \\ &= - \int_{\mathbb{R}} \int_{-\infty}^{\min\{\lambda y + \xi, \bar{x}\} - (\alpha - \beta p)} \varepsilon \varphi(\varepsilon, \xi) d\varepsilon d\xi. \end{aligned}$$

Similarly, the third term simplifies to

$$\begin{aligned} - \frac{\partial}{\partial p} E \left[ \frac{\tilde{q}^2(p, y)}{2\beta} \right] &= - \frac{1}{2\beta} \frac{\partial}{\partial p} \left( \int_{\mathbb{R}} \left[ \int_{-\infty}^{\min\{\lambda y + \xi, \bar{x}\} - (\alpha - \beta p)} (\alpha - \beta p + \varepsilon)^2 \varphi(\varepsilon, \xi) d\varepsilon \right. \right. \\ &\quad \left. \left. + \int_{\min\{\lambda y + \xi, \bar{x}\} - (\alpha - \beta p)}^{\infty} (\min\{\lambda y + \xi, \bar{x}\})^2 \varphi(\varepsilon, \xi) d\varepsilon \right] d\xi \right) \\ &= \int_{\mathbb{R}} \int_{-\infty}^{\min\{\lambda y + \xi, \bar{x}\} - (\alpha - \beta p)} (\alpha - \beta p + \varepsilon) \varphi(\varepsilon, \xi) d\varepsilon d\xi, \end{aligned}$$

so that

$$\begin{aligned} \text{CS}_p(p, y) &= \left(\frac{\alpha}{\beta} - p\right) Q_p - Q - \int_{\mathbb{R}} \int_{-\infty}^{\min\{\lambda y + \xi, \bar{x}\} - (\alpha - \beta p)} \varepsilon \varphi(\varepsilon, \xi) d\varepsilon d\xi \\ &\quad + \int_{\mathbb{R}} \int_{-\infty}^{\min\{\lambda y + \xi, \bar{x}\} - (\alpha - \beta p)} (\alpha - \beta p + \varepsilon) \varphi(\varepsilon, \xi) d\varepsilon d\xi. \end{aligned}$$

Using the expression for  $Q_p$  obtained earlier, we notice that the first term on the right-hand side cancels the last two terms, whence

$$\text{CS}_p(p, y) = -Q(p, y) < 0.$$

Similarly, we obtain for the derivative of the expected consumer surplus with respect to the order-up-to level  $y$  that

$$\begin{aligned} \text{CS}_y(p, y) &= \left(\frac{\alpha}{\beta} - p\right) Q_y + \frac{\lambda}{\beta} \int_{-\infty}^{\bar{x} - \lambda y} \int_{-\infty}^{\min\{\lambda y + \xi, \bar{x}\} - (\alpha - \beta p)} \varepsilon \varphi(\varepsilon, \xi) d\varepsilon d\xi \\ &\quad - \frac{\lambda}{\beta} \int_{-\infty}^{\bar{x} - \lambda y} \int_{\min\{\lambda y + \xi, \bar{x}\} - (\alpha - \beta p)}^{\infty} \min\{\lambda y + \xi, \bar{x}\} \varphi(\varepsilon, \xi) d\varepsilon d\xi, \end{aligned}$$

or, substituting our earlier expression for  $Q_y$ , equivalently,

$$\text{CS}_y(p, y) = \frac{\lambda}{\beta} E \left[ \tilde{d} - \tilde{s} \mid \tilde{s} < \min\{\tilde{d}, \bar{x}\} \right] P(\tilde{s} < \min\{\tilde{d}, \bar{x}\}) > 0.$$

By differentiating  $\text{CS}_p(p, y)$  with respect to the order-up-to level  $y$  we obtain that

$$\text{CS}_{py}(p, y) = -Q_y < 0,$$

since  $Q_y > 0$ . Hence, the expected consumer surplus is submodular with respect to  $(p, y)$ , which completes our proof.  $\blacksquare$

**Proof of Proposition 2.2.** (i) Consider the case of welfare maximization, where  $\omega = 1$ . Then

$$g_{py} = Q_y + pQ_y - L_{py} + \text{CS}_{py},$$

which, using the expression for  $\text{CS}_{py} = -Q_y$ , derived in the proof of Proposition 2.1, implies that

$$g_{yp} = pQ_y - L_{py} < 0,$$

i.e., the intermediary's per-period payoff is submodular in  $(p, y)$ . To establish submodularity of the entire payoff function,  $\Pi_t$ , we need to check that the stochastic state transitions preserve submodularity in  $(p, y)$ , or, equivalently, supermodularity in  $(-p, y)$  or  $(p, -y)$ , in the sense of Smith and McCardle (2002). Specifically we need to show

$$E[g(\tilde{x}_{k-1}(\theta_1))] + E[g(\tilde{x}_{k-1}(\theta_2))] \leq E[g(\tilde{x}_{k-1}(\theta_1 \wedge \theta_2))] + E[g(\tilde{x}_{k-1}(\theta_1 \vee \theta_2))]$$

where  $\theta_1 = (-p_1, y_1)$ ,  $\theta_2 = (-p_2, y_2)$ , and  $\theta_1 \wedge \theta_2$  and  $\theta_1 \vee \theta_2$  denote the meet and join, respectively, of  $\theta_1$  and  $\theta_2$  as defined in a lattice theoretic setting. Recall that the stochastic state transitions in our model are given as

$$\tilde{x}_{k-1} = [s(y) + \xi - d(p) + \epsilon]_+ = [\lambda y + \xi - (\alpha - \beta * p + \epsilon)]_+$$

and, as such, are increasing functions of  $p$  and  $y$ . Then, for any  $\epsilon$  and any  $\xi$ , the stochastic transitions can be seen to preserve the lattice structure, i.e.,  $\tilde{x}_{k-1}$  is a lattice homomorphism. That is, for  $\theta'_1 = \tilde{x}_{k-1}(\theta_1)$  and  $\theta'_2 = \tilde{x}_{k-1}(\theta_2)$ , it is true that  $\tilde{x}_{k-1}(\theta_1 \wedge \theta_2) = (\theta'_1 \wedge \theta'_2)$  and  $\tilde{x}_{k-1}(\theta_1 \vee \theta_2) = (\theta'_1 \vee \theta'_2)$ . Our stochastic state transitions preserve meet and join for any two elements  $\theta'_1$  and  $\theta'_2$ . Hence, we can re-write our earlier expression for the stochastic transitions as

$$E[g(\tilde{x}_{k-1}(\theta_1))] + E[g(\tilde{x}_{k-1}(\theta_2))] \leq E[g(\tilde{x}_{k-1}(\theta_1 \wedge \theta_2))] + E[g(\tilde{x}_{k-1}(\theta_1 \vee \theta_2))]$$

Then, the above holds since, by the supermodularity of  $g$

$$g(\theta'_1) + g(\theta'_2) \leq g(\theta'_1 \wedge \theta'_2) + g(\theta'_1 \vee \theta'_2).$$

This establishes supermodularity of the entire payoff function in  $(-p, y)$  (equivalently, submodularity in  $(p, y)$ ). (ii) Consider now the case of profit maximization, where  $\omega = 0$ . Then the nonnegative term  $Q_y = \lambda P(\tilde{s}_t(y) < \min\{\tilde{d}_t(p), \bar{x}\})$  is not necessarily cancelled out by  $-L_{py} = -\beta\lambda(r+h) \int_{-\infty}^{\bar{x}-\lambda y} \varphi(\lambda y + \xi - (\alpha - \beta p), \xi) d\xi < 0$ , especially if  $\beta$  is small. The proof that supermodularity holds for the entire payoff function follows the same lines as the proof of part (i), where now  $\theta_1 = (p_1, y_1)$  and  $\theta_2 = (p_2, y_2)$ . This completes our proof. ■

**Proof of Proposition 2.3.** The optimal price depends on  $x$  if and only if the optimal order-up-to level is equal to  $x$ . (i) Under welfare maximization, when  $\omega = 1$ , the proof follows immediately from the submodularity of  $g$  of Proposition 2.2 by backward induction, starting for  $t = T$ . (ii) Under profit maximization, when  $\omega = 0$ , part (ii) of Proposition 2.2 states that per-period  $g(p, y)$  can be supermodular, so that for small enough discount factors  $\delta$  the intermediary's total discounted payoff is then also supermodular in  $(p, y)$ , which together with Milgrom and Shannon's (1994) monotonicity theorem concludes our proof. ■

**Proof of Proposition 2.4.** Let  $(p, y), (\hat{p}, \hat{y}) \in \mathbb{R} \times X$  with  $(p, y) \leq (\hat{p}, \hat{y})$ . For  $x = \bar{x}$  it is  $F(x|p, y) = F(x|\hat{p}, \hat{y}) = 1$ . For any  $x \in [0, \bar{x})$  it is

$$F(x|p, y) - F(x|\hat{p}, \hat{y}) = \int_{\mathbb{R}} \left( \int_{\min\{\lambda y + \xi, \bar{x}\} - (\alpha - \beta p) - x}^{\min\{\lambda \hat{y} + \xi, \bar{x}\} - (\alpha - \beta \hat{p}) - x} \varphi(\varepsilon, \xi) d\varepsilon \right) d\xi \geq 0,$$

so that indeed  $F(\cdot|p, y) \preceq_{\text{FOSD}} F(\cdot|\hat{p}, \hat{y})$ . ■

**Proof of Proposition 2.5.** Let  $x \in X$ . The proof proceeds in two steps.

*Step 1.* We show that for any  $y \in [x, \bar{x}]$  the per-period payoff  $g(p, y)$  is coercive in  $p$ . For this, consider

$$\begin{aligned} g_p(p, y) &= Q(p, y) + pQ_p(p, y) - L_p(p, y) + \omega \text{CS}_p(p, y) = \\ &= Q(p, y)(1 - \omega) - \beta \left( p(1 - \omega) + \omega \frac{\alpha}{\beta} + (r + h) \right) P(\tilde{d}_t(p) \leq \tilde{s}_t(y)) + \beta r - \\ &\omega \left( \int_{\mathbb{R}} \int_{-\infty}^{\min\{\lambda y + \xi, \bar{x}\} - (\alpha - \beta p)} \varepsilon \varphi(\varepsilon, \xi) d\varepsilon d\xi - \int_{\mathbb{R}} \int_{-\infty}^{\min\{\lambda y + \xi, \bar{x}\} - (\alpha - \beta p)} (\alpha - \beta p + \varepsilon) \varphi(\varepsilon, \xi) d\varepsilon d\xi \right) \\ &= Q(p, y)(1 - \omega) - \beta \left( p(1 - \omega) + \omega \frac{\alpha}{\beta} + (r + h) \right) P(\tilde{d}_t(p) \leq \tilde{s}_t(y)) + \beta r - \omega \zeta \end{aligned}$$

where  $\zeta = \left( E[\tilde{\varepsilon}|\tilde{d}_t(p) < \tilde{s}_t(y)] - E[\tilde{q}|\tilde{d}_t(p) < \tilde{s}_t(y)] \right) P(\tilde{d}_t(p) < \tilde{s}_t(y))$ . Since

$$\lim_{p \rightarrow \infty} P(\tilde{d}_t(p) \leq \tilde{s}_t(y)) = \lim_{p \rightarrow \infty} P(\tilde{d}_t(p) \leq \tilde{s}_t(\bar{x})) = 1,$$

for any  $\eta \in (0, 1)$  there is a  $\hat{p}_\eta$ , independent of  $y$ , such that  $p \geq \hat{p}_\eta$  implies that  $P(\tilde{d}_t(p) \leq \tilde{s}_t(y)) > 1 - \eta$ . Hence, by setting  $\eta = 1/2$  for any  $\rho > 0$ , we have the following:

$$g_p(p, y) \leq \bar{x} - \beta \left( p(1 - \omega) + \omega \frac{\alpha}{\beta} + (r + h) \right) (1 - \eta) + \beta r - \omega \zeta < -\rho\beta/2 < 0 \quad (2.2)$$

for all  $p \geq \max\{\hat{p}(x), r + h - \omega \left( \frac{\alpha}{\beta} - \frac{2\zeta}{\beta} \right) + \rho + \frac{2\bar{x}}{\beta}\}$ , where  $\hat{p}(x)$  is such that  $P(\tilde{d}_t(\hat{p}(x)) \leq \tilde{s}_t(x)) \geq 1/2$ . The upper bound  $\bar{p}(x)$  obtains by taking the limit for  $\rho \rightarrow 0^+$ . The upper bound  $\bar{p}(x)$  is nonincreasing in the state  $x$ , since  $P(\tilde{d}_t(p) \leq \tilde{s}_t(x))$  is nondecreasing in  $x$  for any given  $p$ . We therefore conclude that increasing  $p$  beyond  $\bar{p}(x)$  can only decrease expected payoffs. Let us now consider  $p = 0$ . For any  $y \in [x, \bar{x}]$  it is

$$g_p(0, y) = Q(0, y)(1 - \omega) - \beta \left( \omega \frac{\alpha + \tilde{\varepsilon}}{\beta} + (r + h) \right) P(\tilde{d}_t(p) \leq \tilde{s}_t(y)) + \beta r + \omega \zeta > 0,$$

so that a price of zero cannot be optimal: increasing  $p$  yields a strict improvement in expected payoffs. We have thus shown that at time  $t = T$ , any price  $p_T$  that maximizes per-period rewards must lie in  $[0, \bar{p}(x)]$ .

*Step 2.* We now show by induction that for any  $t \in \{1, \dots, T-1\}$  the expected value at time  $t+1$  is strictly decreasing in  $p$  if only  $p$  is large enough. Assume that this is true for  $\bar{\Pi}_{t+1}(p, y)$ . Then

$$\bar{\Pi}_t(p, y) = g(p, y) + cE[\tilde{x}_{t+1}|p, y] + \delta\bar{\Pi}_{t+1}(p, y)$$

is decreasing in  $p$  for large enough  $p$  if  $g(p, y) + cE[\tilde{x}_{t+1}|p, y]$  is decreasing in  $p$  for large enough  $p$ . We have that

$$\frac{\partial}{\partial p}E[\tilde{x}_{t+1}|p, y] = \beta P(\tilde{d}_t(p) \leq \tilde{s}_t(y)) \leq \beta,$$

so that we can conclude from (2.2) with  $\rho = (2 + \hat{\rho})c > 0$  for any  $\hat{\rho} > 0$  that

$$g_p(p, y) + \frac{\partial}{\partial p}E[\tilde{x}_{t+1}|p, y] \leq c\beta - c(2 + \hat{\rho})\beta/2 = -\hat{\rho}\beta/2 < 0$$

for all  $p \geq \max\{\hat{p}(x), r + h - \omega\left(\frac{\alpha}{\beta} + \frac{2\zeta}{\beta}\right) + \rho + \frac{2\bar{x}}{\beta}\}$ , where  $\hat{p}(x)$  is such that  $P(\tilde{d}_t(\hat{p}(x)) \leq \tilde{s}_t(x)) \geq 1/2$ . As in Step 1, the upper bound  $\bar{p}(x)$  obtains by taking the limit for  $\hat{\rho} \rightarrow 0^+$ , which concludes our proof.  $\blacksquare$

**Proof of Proposition 2.6.** (i) The existence of an optimal order level is guaranteed by the Weirstrass theorem, which states that any continuous real-valued function on a non-empty compact metric space will obtain a maximum value. Hence by the continuity of the function  $\Pi_t$  and the compactness of  $y$ , i.e.,  $y \in \{x, x+z\}$ , we are guaranteed the existence of  $y^*$ . The compactness requirement is not immediately satisfied for  $p$  but is established in Proposition 2.5.

The proof of  $y^*$  increasing in  $x$  is a proof by contradiction. Let  $y^*$  denote the optimal  $y$  for some state  $x$  and let  $\hat{y}^*$  denote the optimal  $y$  in some other state  $\hat{x}$ , where,  $\hat{x} > x$ . Assume that  $\hat{y}^* < y^*$ . We then have  $\hat{y}^* \in \{x, x+z\}$  and  $y^* \in \{x', x'+z\}$ , since it is  $x < x' < \hat{y}^* < y^* \leq x+z < x'+z$ , which establishes the feasibility of  $\hat{y}^*$  in state  $x$  and, likewise, the feasibility of  $y^*$  in state  $x'$ . By the definition of the optimum we have, in state  $x$ ,  $y^* > \hat{y}^*$ , but, in state  $x'$  it must be  $\hat{y}^* > y^*$ . We have reached our contradiction, and we conclude that  $y^*$  is increasing in  $x$ . Note that  $y$  is set-valued and that in the case of multiple maxima the above argument will hold for any given maxima, i.e., for any element  $y^*$  in the set of maxima.

That  $y^*$  is also increasing in  $z$  is readily seen: an increase in  $z$  only serves to *increase* the range of possible values for  $y^*$ , so that any  $y^*$  feasible for  $z$  is also feasible for  $z'$ , where  $z' > z$  and, hence,  $y^*$  cannot be decreasing in  $z$ .

Under the conditions established in Proposition 2.3,  $p^*$  is decreasing in  $x$  and  $z$  for cases of submodular  $g(p, y)$  and increasing in  $x$  and in  $z$  for cases of supermodular  $g(p, y)$ . The extension of these properties of  $g(p, y)$  to the multi-period (dynamic) model, and hence the general result, requires that the stochastic transitions satisfy sub- or supermodularity; this is verified in Proposition 2.2 above.

(ii) Let  $(p_t^*, y_t^*)$  be the optimal decision pair for the unconstrained problem. For  $y_t^* \in \{x, x + z\}$ , this is still the optimal pair. If  $y_t^* < x$ , it is optimal to choose  $y_t^* = x$ . To see that this is optimal, consider choosing an alternate decision pair  $(x', p')$ , where  $x' > x$ . By concavity, we must have, for the pair  $(x, p'')$  on the line connecting  $(p_t^*, y_t^*)$  and  $(x', p')$  that  $\Pi_t(x, p'') \leq \Pi_t(x', p')$ , which contradicts the optimality of  $x' > x$ . A similar argument holds in the other direction. If  $y_t^* > x + z$ , it is optimal to choose  $y_t^* = x + z$ . If we choose instead the pair  $(\hat{x}, \hat{p})$  where  $\hat{x} < x + z$ , then we will arrive at the contradiction  $\Pi_t(\hat{x}, \hat{p}) \leq \Pi_t(x + z, \hat{p})$ . This, combined with the submodularity of  $\Pi_t$ , established in Proposition 2.3 ensures that  $p^*$  is decreasing in the inventory level,  $x$ . These two conditions - decreasing  $p^*(x)$  and optimality of  $x$ , in the case that  $y^* < x$ , or, if  $y^* > x + z$ , optimality of  $x + z$  are precisely what's needed to satisfy the definition of a base-stock-list-price policy. ■

**Proof of Proposition 2.8.** The single-period value function  $g(p, y)$  exhibits increasing differences in  $\omega$  and  $(p, y)$ . To see this, recall  $CS_p = \omega(-Q(p, y)) \leq 0$ , hence  $CS_{-p} = \omega(Q(p, y)) \geq 0$ . Also,  $CS_y \geq 0$ . Then clearly for  $\omega \geq 0$ , both  $CS_{-p\omega}$  and  $CS_{y\omega}$  are positive. Then from Proposition 2.1, we know that  $g(p, y)$  is supermodular in  $(-p, y)$ . The result follows from Milgrom and Shannon's (1994) Monotonicity Theorem. We note that since the stochastic transitions are unaffected by  $\omega$ , we do not have to verify that they preserve this property. ■

**Proof of Proposition 2.9.** The proof is straightforward. We want to show that  $E[v(p, y; C_1)] \geq E[v(p, y; C_2)]$ , where  $C_1 \leq C_2$  is some cost parameter and  $v(p, y)$  is the primitive value function. The above expression holds by the linearity of the expectation operator and the fact that  $v(p, y; C_1) \geq v(p, y; C_2)$  for all  $C_1 \leq C_2$ , where

$$\begin{aligned} v(p, y) = & p \min\{\alpha - \beta p + \epsilon, \min\{\lambda y, \bar{x}\}\} - c(y - x) - \\ & \max\{h(\min\{\lambda y, \bar{x}\} - (\alpha - \beta p + \epsilon)), r(\alpha - \beta p + \epsilon - \min\{\lambda y, \bar{x}\})\} + \\ & \omega(\hat{p} - p) \min\{\alpha - \beta p + \epsilon, \min\{\lambda y, \bar{x}\}\} + (\min\{\alpha - \beta p + \epsilon, \min\{\lambda y, \bar{x}\}\})^2 / 2\beta. \end{aligned}$$

is monotone decreasing in all costs, i.e., in  $c$ ,  $h$ , and  $r$ . Unlike the Proof of 2.1 we do *not* need to verify that the property  $P$  "decreasing" is preserved by the stochastic transitions,



since the costs do not effect these transitions. ■

**Proof of Proposition 2.10.** (i) From Athey (2002) we know that log-supermodularity of a primitive value function is preserved by integration with a log-supermodular density function, i.e., the expected value will be log-supermodular if both the primitive function and the density are log-supermodular. Furthermore, log-supermodularity of the expected reward function is both necessary and sufficient to conclude that the argmax is increasing in a given parameter of the density. To be precise, from Athey (2002) we have that for log-supermodular  $u(\mathbf{x}, \mathbf{s})$  and log-supermodular density  $f(\mathbf{s}; \theta)$ ,  $\arg \max_{x \in B} \int u(\mathbf{x}, \mathbf{s}) f(\mathbf{s}; \theta) d\mu(\mathbf{s}) \uparrow$  in  $\theta$  and  $B$ .

The extension of Athey's (2002) results to a dynamic setting and a treatment of monotone comparative statics for dynamic programs comes from Smith and McCardle (2002). For a given property  $P$  to be preserved in a dynamic programming model, we require, additionally, that the stochastic transitions satisfy the property  $P$  in the sense of stochastic dominance. Specifically, we will need to establish that log-supermodularity is preserved by the stochastic transitions in the sense of Smith and McCardle (2002). We will proceed in a manner similar to that in Proof 2.1 above, where we demonstrated that supermodularity is preserved by the stochastic transitions in our model.

From Milgrom and Weber (1982) a density is log-supermodular if and only if it satisfies the monotone likelihood ratio order (MLR). A density function satisfies MLR if  $\frac{f(x; \theta_1)}{f(x; \theta_2)}$  is nondecreasing in  $x$  for  $\theta_1 > \theta_2$ , as discussed in Karlin and Rubin (1956). It is well known that the univariate normal distribution satisfies MLR in its mean (see e.g., Karlin and Rubin (1956). That is,  $\varphi(\epsilon; \mu_\epsilon)$  is log-supermodular. The extension of this property to the multivariate normal, as discussed in Karlin and Rinott (1980), requires that the following technical condition be met: for the matrix  $B$ , where  $\Sigma^{-1} = B = \|b_{ij}\|_{i,j=1}^n$ , we must have  $b_{ij} \leq 0$  for all  $i \neq j$ . We note that  $\Sigma^{-1}$  denotes the inverse of the covariance matrix  $\Sigma$ .

For the bivariate normal, this condition is satisfied for any  $\rho \geq 0$ , where  $\rho$  is the correlation coefficient between the two variables. Hence, for the case of positive correlation, our joint density  $\varphi(\epsilon, \xi; \mu_\epsilon)$  is log-supermodular.

The proof proceeds in two steps. In Step 1, we verify the log-supermodularity of our primitive function, which, coupled with the log-supermodularity of the density function, implies log-supermodularity of the per-period payoff function. In Step 2, we verify that the stochastic transitions do indeed preserve the property  $P$  of interest, log-supermodularity.

*Step 1.*

Our primitive value function (Athey's  $u(\mathbf{x}, \mathbf{s})$  above) is given as follows for the case of pure

demand uncertainty:

$$\begin{aligned} v(p, y, \epsilon) &= p \min\{\alpha - \beta p + \epsilon, \min\{\lambda y, \bar{x}\}\} - c(y - x) - \\ &\max\{h(\min\{\lambda y, \bar{x}\} - (\alpha - \beta p + \epsilon)), r(\alpha - \beta p + \epsilon - \min\{\lambda y, \bar{x}\})\} + \\ &\omega(\hat{p} - p) \min\{\alpha - \beta p + \epsilon, \min\{\lambda y, \bar{x}\}\} + (\min\{\alpha - \beta p + \epsilon, \min\{\lambda y, \bar{x}\}\})^2/2\beta. \end{aligned}$$

Recall that a function  $v(p, y)$  is log-supermodular if  $\frac{\partial^2}{\partial p \partial y} \log v(p, y) \geq 0$ . Specifically, in a multivariate setting, it must be  $\frac{\partial^2}{\partial x_i \partial x_j} \log v(\mathbf{x}) \geq 0$  for all  $i \neq j$ .

We will check that  $\frac{\partial^2}{\partial p \partial y} \log v(p, y, \epsilon) \geq 0$ ,  $\frac{\partial^2}{\partial p \partial \epsilon} \log v(p, y, \epsilon) \geq 0$ , and  $\frac{\partial^2}{\partial y \partial \epsilon} \log v(p, y, \epsilon) \geq 0$  under both eventualities:  $\tilde{d} \geq s(y)$  and  $s(y) > \tilde{d}(p)$ .

Under profit maximization, i.e.,  $\omega = 1$ , we have

*Case 1.*  $\tilde{d}(p) \geq s(y)$ :

$$\begin{aligned} \frac{\partial^2}{\partial p \partial y} \log v(p, y, \epsilon) &= \frac{\lambda h}{v^2} [\tilde{d}(p) - \beta(p + h)] \geq 0 \\ \frac{\partial^2}{\partial p \partial \epsilon} \log v(p, y, \epsilon) &= \frac{v - (p + h + \epsilon)[\tilde{d}(p) - \beta(p + h)]}{v^2} \geq 0 \\ \frac{\partial^2}{\partial y \partial \epsilon} \log v(p, y, \epsilon) &= \frac{\lambda h}{v^2} (p + h\epsilon) \geq 0 \end{aligned}$$

*Case 2.*  $s(y) > \tilde{d}(p)$ :

$$\begin{aligned} \frac{\partial^2}{\partial p \partial y} \log v(p, y, \epsilon) &= \frac{\lambda(v - (p + r))[\tilde{s}(y) + r\beta]}{v^2} \geq 0 \\ \frac{\partial^2}{\partial p \partial \epsilon} \log v(p, y, \epsilon) &= \frac{r}{v^2} [\tilde{s}(y) + r\beta] \geq 0 \\ \frac{\partial^2}{\partial y \partial \epsilon} \log v(p, y, \epsilon) &= \frac{r}{v^2} (p\lambda + r\lambda) \geq 0 \end{aligned}$$

This completes the first step.

*Step 2.*

Now we need to verify that log-supermodularity is preserved by our stochastic transitions in the sense of Smith and McCardle (2002). Using the fact that a function  $g$  is log-supermodular if and only if the function  $f$  is supermodular, where  $f = \log g$ , we will check that

$$E[f(\tilde{x}_{k-1}(\theta_1))] + E[f(\tilde{x}_{k-1}(\theta_2))] \leq E[f(\tilde{x}_{k-1}(\theta_1 \wedge \theta_2))] + E[f(\tilde{x}_{k-1}(\theta_1 \vee \theta_2))]$$

where  $f = \log g(p, y)$  and  $g(p, y)$  is our log-supermodular per-period payoff function. Since  $f$  is supermodular, the proof is the same as that in Proof 2.1 above: supermodularity is preserved by the stochastic transitions. We conclude that since  $\log \Pi(x_t)$  is supermodular, from

Topkis (1978)  $\Pi(x_t)$  is log-supermodular. We are done with part (i). Part (ii) follows immediately, since the assumptions guarantee that the conditions verified above do indeed hold. ■

## Appendix B: Operational Information for MWD

Table 2.5 below details MWD's rationing policy under the Incremental Interruption and Conservation Program (IICP), which has since been replaced by the Water Surplus and Drought Management Plan (WSDM), adopted in 1999. Stage I of the IICP was adopted in December of 1990. Then, in January of 1991, a Stage III policy was put in place. In February 1991, MWD declared a Stage IV, which remained in place until March 1992, when it downgraded to a Stage III, finally returning to Stage I in April of 1992.

Stage	Reduction in Non-Firm Deliveries	Conservation Goal for Firm Deliveries
Stage I	Voluntary	Goal of 10%
Stage II	20%	5%
Stage III	30%	10%
Stage IV	40%	15%
Stage V	50%	20%

Table 2.3: MWD's Incremental Interruption and Conservation Plan, Stages I-IV

In 2002 MWD moved to an “unbundled” rate system, whereby rates are presented as the sum of the costs of providing different services. (Previous to 2004 MWD had a lumped rate, setting a price for, say, “full service (non-interruptible) treated” water or “untreated interim agricultural” water, without delineating the different cost components of providing this water.) Tier 1 Supply Rate represents the current cost of acquiring one unit (one af) of water; the Tier 2 Supply Rate represents the cost of augmenting the existing water supply, by an af. The System Access Rate is the cost of infrastructure for water deliveries. The Water Stewardship Rate goes towards environmental protection. The Treatment Surcharge is the cost of treating an af of water. Finally, the last two charges are fixed fees, where the Readiness-to-Serve Charge (RTS) is designed to cover the cost of providing emergency storage services, in case of supply disruption in the event of an earthquake or of a severe drought. The \$80M fee is divided amongst the 26 member agencies based on the rolling 10-year average of deliveries to each agency; that is, if an agency's 10-year average of deliveries

represents 10% of the total average 10-year deliveries, then that agency will cover 10% of the RTS. Finally, the Capacity Charge for each agency is levied on the maximum daily flow to the agency within the last three years, between May 1 and September 20. This charge is designed such that it “recovers the cost of providing peak capacity within the distribution system.” In our model, we consider only the price full service treated water (Tier 1 and Tier 2 water with all of the additional charges, i.e., with the system access rate, water stewardship rate, system power rate, and treatment surcharge added to the Tier 1 and Tier 2 base fee.)

Description of Charges	Rate for FY 2006
Tier 1 Supply Rate (\$/af)	\$73
Tier 2 Supply Rate (\$/af)	\$169
System Access Rate (\$/af)	\$152
Water Stewardship Rate (\$/af)	\$25
System Power Rate (\$/af)	\$81
Treatment Surcharge (\$/af)	\$122
Readiness-to-Serve Charge (\$M)	\$80M
Capacity Charge (\$/cfs)	\$6,800

Table 2.4: Unbundled Water Rates for Financial Year 2006

Table 2.5 reports the rates MWD has charged over the last twenty one years – through the end of 2006 – for full service treated water, which is what we model. As noted in Table 2.5 above, MWD offers different types of water, at different rates, including untreated full service water, treated and untreated “interim agricultural” water, and treated and untreated “seasonal storage” water. The full service water represents the largest portion of MWD deliveries and is considered “non-interruptible,” whereas the interim agricultural and seasonal storage supplies are deemed “interruptible.” In 2003, MWD introduced a block pricing system, with “Tier 1” and “Tier 2” charges for water consumed above a certain level. That is, all member agencies were provided the opportunity to enter into long-term (10-year) contracts with MWD for delivery of up to 90% of their historical water use at the Tier 1 rate. Usage above this amount would be charged the Tier 2 rate. The Tier 2 rate is denoted in parentheses, after the Tier 1 rate, for years 2003–2006 above.

Table 2.5 documents the estimated shortage costs by sector (per af of undelivered water), from the CIC Research, Inc., study, “The Economic Impact on San Diego County of Three Levels of Water Delivery: 80, 60, or 40 Percent Occurring for Two Months or Six Months.” The total costs per undelivered af of water sum to \$5,885. One af cut in supplies is assumed to be uniformly applied across all sectors. (“F.I.R.E” stands for “Financial Services and Real

Year(s)	Full Service Treated Rate
1986-1990	\$230/af
1991	\$261/af
1992	\$322/af
1993	\$385/af
1994	\$412/af
1995-1996	\$426/af
1997-2002	\$431/af
2003	\$408/af (\$489/af)
2004	\$418/af (\$499/af)
2005	\$443/af (\$524/af)
2006	\$453/af (\$549/af)

Table 2.5: Historical Water Rates: 1986-2006

Estate," "Trans. and P.U." is Transportation and Public Utilities, and "O.R.P." is Other Resource Production.)

Sector	Shortage Cost
Agriculture	\$569/af
O.R.P.	\$123/af
Construction	\$449/af
Trans. & P.U.	\$255/af
Manufacturing	\$763/af
Wholesale Trade	\$149/af
F.I.R.E.	\$1,082/af
Retail Trade	\$575/af
Services	\$1,858/af
Government	\$61.3/af
Total	≈ \$5,885/af

Table 2.6: Shortage Cost Estimates by Sector

# Chapter 3

## Bilateral Option Contracting

### 3.1 Motivation

Option contracts are widely used in the electricity sector as a means of coordinating stochastic demand and supply and thereby avoiding costly shortfalls. These contracts could serve a similar purpose in the water sector, where low water supply and high demand years are generally coupled. Over the past several years, option contracts have been signed in the California water market between the state's major intermediary, the Metropolitan Water District (MWD) of Southern California, and Sacramento Valley farmers – and more recently between the San Diego County Water Authority (SDCWA) and several irrigation districts.

The introduction of option contracts in the water sector offers two primary advantages: the first is the ability to eliminate downside risk by ensuring that excess demand can be met (by calling the options), also a chief concern in the electricity sector. The second is unique to the water sector, namely the feasible reallocation of water resources from agricultural to urban uses. As discussed in Chapter 1, institutional impediments to establishing a water market in California are largely attributable to the focus on permanent water rights transfers. Market friction in the form of high transaction cost and institutional resistance has impeded past transfers. Temporary water transfers, as achieved under option contracts (and forward contracts), reduce the transaction cost and institutional resistance associated with permanent transfers, thereby facilitating trade. As such, they offer the potential to ultimately achieve a market solution to the water reallocation problem.

The contracts signed to date in the water market are principally bilateral contracts, negotiated between two parties. In the bilateral contracting model developed in this chapter,

we assign the price-setting power to the seller (designating him as a von Stackelberg leader in the contracting interaction). This assumption is not entirely representative, given the common practice of open-ended negotiation between the two contracting parties. It does, however, provide insight into contract structure where the bargaining power of the buyer is limited.<sup>1</sup> As the number of buyers in the market increases, and there is additional information dissemination regarding buyers actual shortage costs, the seller is likely to appropriate more of the price-setting power. The model thus gains predictive power and also provides a benchmark for sellers in terms of seller-optimal pricing rules.

Among the questions we set out to answer are the following: (1) what is the structure of the optimal option contract, i.e., what are the quantities and prices associated with the optimal contract? (2) how does the presence of uncertainty regarding the buyer's downstream demand and the price of the outside (market) option impact the structure of the contract? (3) how does the contract change when the buyer has an outside supply option, which could include access to a spot market, and (4) how does the allocation achieved under bilateral contracting compare to the socially optimal allocation?

Regarding the latter, we expect there to be deadweight losses in a system where there is a single seller. Indeed, this is the case. The single seller exhibits monopolistic behavior, and, as we demonstrate, our option-pricing results can be interpreted in light of the standard theory of multi-product monopoly pricing of complementary goods (where the option is a perfect complement to the strike). (See e.g., Andriychenko et al. (2006) and Economides et al. (2006).)

In a previous paper on option contracts for capital-intensive goods, Wu et al. (2002) present a bilateral contracting model with joint spot market access and certain demand. They demonstrate that the optimal option price is set according to the standard monopoly-pricing rule, the inverse-elasticity rule (although they do not explicitly identify it as such). The strike price is set equal to the seller's marginal cost of an exercised option, i.e., to the delivery cost of the option. We extend the analysis by Wu et al. to the case where the buyer faces uncertain downstream demand. We also explicitly consider the dependence of the contracting outcome on the reserve value of the buyer (not just the seller).<sup>2</sup> The introduction of uncertain demand

---

<sup>1</sup>In the existing water market, the buyers (urban water districts) can, for instance, leverage political power to pressure sellers.

<sup>2</sup>Note: Wu et al.'s model introduces a "risk factor." The risk factor in their model is a constant that represents, alternately, the percentage of supply that the seller can expect to offload on the market or the probability of finding a buyer on the given day. A risk factor of one corresponds to the "no risk" case, and the seller can offload his supply with certainty in the second period, in which case Wu et al. find that there is no incentive to contract. The problem with this risk factor is its assumed independence from the the

not only changes the price structure of the contract, it also changes the conditions under which contracting arises. Wu et al. find that with joint spot market access contracting is only feasible when one party faces a risk, or loss, from not contracting, i.e., the seller may face some positive probability of not being able to unload his supply onto the market. In the presence of uncertain downstream demand, we find more general conditions for contracting: if the marginal value of avoided shortage is high enough, then contracting occurs *irrespective of the loss factor*.

The rest of the chapter is organized as follows: in Section 3.2 we introduce the general bilateral contracting model. Before proceeding with the analysis of the model, we discuss first-best outcomes, in Section 3.3. We then discuss conditions under which contracting is feasible, or desirable, in Section 3.4. In Sections 3.5-3.7 we discuss the structure of the optimal contract under three differing assumptions with regards to the buyer's reserve value. First, we assume that the buyer has no outside option for supply. We then relax this assumption and assume the presence of an outside option at a fixed price. Finally, in Section 3.7 we return to the setting of Wu et al., namely joint spot market access, and examine the optimal contract with the inclusion of demand uncertainty. We conclude with an application of the model to contracting in the California water market, in Section 3.8, and a summary of the chapter in Section 3.9. Proofs of many of the propositions can be found in the Appendix.

## 3.2 The Model

We develop a general two-stage model of bilateral contracting for options, where the buyer may or may not have access to an outside supply option. When the buyer has an outside supply option, it may be of a known (fixed) price, or the price may remain uncertain. If we assume that the buyer and seller both have access to a spot market for water (as we will do in Section 3.7), then the buyer's outside supply option is to purchase from the spot market in the second period, at a price that remains uncertain in the first period.

In the first period, the buyer and seller contract, wherein a contract size, or number of options,  $Q$  is specified as well as an upfront fee  $p$  and a strike price  $s$  per option. In the second period, the uncertainty (present in the first period) is resolved, the buyer calls his options, goods are delivered, and the seller (and buyer, if he has access) may also transact on the spot market.

---

quantity of options being traded. In reality, we would expect the risk of, say, not being able to offload one's supply in period two to depend directly on the total volume of supply being offloaded.



The buyer's problem is to decide a contract level  $Q$ . The seller's problem is to determine what price  $p$  to charge per option and what strike price  $s$  to offer, i.e., to set  $(p, s)$ . As in Wu et al. (2002), we assume that the buyer and seller are playing a von Stackelberg game with the seller as the leader. Our analysis assumes perfect information – the seller knows a complete description of the buyer's value function. More precisely, he knows exactly how many options the buyer will demand at a given price pair  $(p, s)$ .

The buyer and the seller both face uncertainty. However, unlike in Wu et al. (2002) who assume a single source of uncertainty – the future spot market price – we introduce an additional source of uncertainty for the buyer. The buyer is an intermediary and faces uncertainty regarding his future downstream demand.

The seller faces uncertainty regarding the spot market price  $\tilde{\mu}$  at which he can offload his supply. The spot market will have two alternate interpretations in our model. In the case where the buyer and the seller have joint spot market access, the spot market is for the good being contracted. For example, in Section 3.7 where there is joint market access, we are modelling contracting for water in the presence of a spot market for water. (Although such a spot market does not currently exist, it may in the future.) The alternative interpretation of the spot market – the interpretation when the buyer does not have access to it – is a market for a good *for which the contracted good is a production input*. In the existing California water market, this is the representative case: the seller (a farmer) faces a spot market for a crop produced with water (the contracted good) as an input.<sup>3</sup> For example, an option contract may arise between a rice farmer (seller) and a municipal water district (buyer). There is no spot market for water. However, there is a spot market for rice.

The buyer's uncertain downstream demand  $\tilde{D}$  has density  $f_D(D)$ . The uncertain spot price  $\tilde{\mu}$  faced by the seller has density  $f_\mu(\mu)$ . Here  $F_D$  will denote the cumulative distribution function (cdf) of demand, and  $F_\mu$  will denote the cdf of the spot price. We assume independence of these two sources of uncertainty.<sup>4</sup> The mean of the spot market price faced by the seller is  $\bar{m}$ . The (possibly) uncertain price of the buyer's outside supply option  $\tilde{\xi}$  has cdf  $F_\xi$ . When the buyer has access to the same spot market for the good as the seller  $\tilde{\xi} = \tilde{\mu}$  and both the seller and the buyer face the same distribution of outcomes for the good's second-period market price.

---

<sup>3</sup>For simplicity, we will assume in our model that the seller can produce exactly one unit of this crop per unit of available water and, hence, has exactly one unit to sell on the spot market per unit of unclaimed water, i.e., the production function is unity.

<sup>4</sup>This assumption seems highly reasonable when the spot market in question is for, say, a global agricultural commodity. However, when the spot market is for water, the assumption is weaker. We would expect to generally find a higher spot market price for water in periods of high demand.

We include one additional uncertainty – that surrounding the possibility of an infrastructure failure. The infrastructure failure could be due to a physical collapse of infrastructure or due to congestion. We represent the possibility of an infrastructure failure as the random variable  $\tilde{\rho}$ , assumed to be uniformly distributed on  $(0, 1)$ , and hence with mean  $\rho = 0.5$ .

In the second-period, the buyer solves the following maximization problem

$$\max_{\hat{z}, z \leq Q} \{V(z, \hat{z}; D) - sz - (\xi + \tau)\hat{z}\}, \quad (3.1)$$

where  $V(z, \hat{z}; D)$  is the buyer's gross payoff when delivering the amount  $z + \hat{z}$  to downstream users, given realized demand  $D$ ,  $\tau$  is the “market friction” coefficient or transaction cost associated with the buyer's outside supply option, and  $\xi$  is the realized price of the outside supply. Then  $z$  and  $\hat{z}$  are the amount of supply that the buyer gets from exercising his options and from direct purchase from the market (or some alternate source of supply), respectively. We have

$$z(Q, s; D) = \begin{cases} g(s, D)_+, & \text{if } D < \bar{D} \text{ and } s \leq \xi + \tau \\ Q, & \text{if } D \geq \bar{D} \text{ and } s \leq \xi + \tau, \\ 0, & \text{otherwise,} \end{cases}$$

and

$$\hat{z}(Q, \tau, \xi; D) = \begin{cases} [g(\xi + \tau, D) - Q]_+, & \text{if } s \leq \xi + \tau \\ g(\xi + \tau, D)_+, & \text{otherwise.} \end{cases}$$

where  $g(s, D) = \{z : V_z(z, D) = s\}$  and  $D(s) = \{D : g(s, D) = 0\}$  and  $\bar{D}(s, Q) = \{D : g(s, D) = Q\}$ .

At a price pair  $(p, s)$  the buyer's first-period problem is to maximize his expected payoff

$$\max_{Q \geq 0} \{U(Q, s, \tau) - pQ\}, \quad (3.2)$$

which is given as follows:

$$U(Q, s, \tau) - pQ = \rho\Phi_1 + (1 - \rho)\Phi_2 - pQ$$

where

$$\begin{aligned} \Phi_1 = & \int_0^{s-\tau} \int_0^{D(\xi+\tau)} V(0, D) dF(D) dF(\xi) + \\ & \int_0^{s-\tau} \int_{D(\xi+\tau)}^{\kappa} [V(g(\xi + \tau, D), D) - (\xi + \tau)g(\xi + \tau, D)] dF(D) dF(\xi) \\ & + (P(\tilde{\xi} \geq s - \tau) \left[ \int_0^{D(s)} V(0, D) dF(D) dF(\xi) + \int_{D(s)}^{\bar{D}(s, Q)} [V(g(s, D), D) - sg(s, D)] dF(D) \right] \end{aligned}$$

$$\begin{aligned}
& + \int_{s-\tau}^{\gamma} \int_{\bar{D}(s,Q)}^{\bar{D}(\xi+\tau,Q)} [V(Q, D) - sQ] dF(D) dF(\xi) \\
& + \int_{s-\tau}^{\gamma} \int_{\bar{D}(\xi+\tau,Q)}^{\kappa} [V(g(\xi + \tau, D), D) - sQ - (\xi + \tau)(g(\xi + \tau, D) - Q)] dF(D) dF(\xi) - pQ
\end{aligned}$$

and

$$\begin{aligned}
\Phi_2 = & \int_0^{\gamma} \int_0^{D(\xi+\tau)} V(0, D) f_D(D) f_{\xi}(\xi) dD d\xi + \\
& \int_0^{\gamma} \int_{D(\xi+\tau)}^{\kappa} [V(g(\xi + \tau, D), D) - (\xi + \tau)g(\xi + \tau, D)] f_D(D) f_{\xi}(\xi) dD d\xi
\end{aligned}$$

where  $\kappa$  is the upper bound of the support of the demand distribution,  $\gamma$  is the upper bound of the support of the price of the outside supply option  $\tilde{\xi}$ . The seller's profit-maximizing objective function is given as follows:

$$\max_{p,s} \{pQ(p, s) + (s - b)y + E(\tilde{m} - b)(K - y)\}$$

where  $\tilde{m}$  is the random spot market price,  $b$  is the marginal cost of delivery of the good,  $K$  is the seller's total capacity, and  $y$  is the expected number of options to be executed given that  $Q$  are purchased:

$$y = P(\tilde{\xi} \geq s - \tau) [Q(1 - F_D) + \bar{g}]$$

where  $\bar{g} = \int_{D(s)}^{\bar{D}(s,Q)} g(s, D) dF(D)$ .

We observe that  $y$  is proportional to the probability that the price of the buyer's outside supply option  $\tilde{\xi}$  exceeds the strike price  $s$  minus the transaction fee  $\tau$ . In the case where the outside supply option has a fixed price  $\xi$ ,  $P(\tilde{\xi} \geq s - \tau)$  collapses to 1, for  $s < \mu - \tau$  (in which case  $y = [Q(1 - F_D) + \bar{g}]$ ), or 0, for  $s > \mu + \tau$  (in which case  $y = 0$ ).

We note that the buyer expects to exercise all  $Q$  options that he purchases (equivalently, he purchases only options that he expects he will use). In contrast, the number of options that the seller expects to be executed *given that  $Q$  are purchased* is less than  $Q$ .

### 3.3 First-Best Allocation

Under bilateral option contracting we do not generally achieve the first-best allocation. There are two potential sources of inefficiency. First, an upfront fee is inefficient. As we will show, under joint surplus maximization the optimal allocation is that which sets an option price  $p$

of zero.<sup>5</sup> This allows maximum flexibility in the contract, with the buyer electing to hold as many options as he would conceivably exercise. In general, a profit-maximizing seller will extract an upfront fee, generating the first source of inefficiency: the buyer holds too few options from a social welfare standpoint, e.g., less than he may need in the second period.

The second source of inefficiency is the biasing of the optimal second-period decision. An optimal allocation places the resource in question in the hands of the user who values it most highly at that point in time. Who values the resource most highly (the seller or the buyer) is uncertain in the first period, and, hence, setting a rigid strike price can create ex-post inefficient outcomes. If the contract can be renegotiated at the time of exercise, this inefficiency is eliminated. For instance, in financial markets, physical delivery of goods under contract is rarely taken. In the second period, the good under contract is sold on the market to the highest bidder and money exchanges hands between the contract seller and buyer. If the contract must be taken on delivery, inefficiencies can arise when the spot market price and the strike price of the contract do not coincide.

As we show below, the social-welfare maximizing option contract sets prices  $(0, \bar{m})$ , where  $\bar{m}$  is the expected spot market price. This contract eliminates the inefficiency associated with an upfront fee. However, it fails to achieve first-best for the reason just given – the rigid strike price biases the second-period allocation decision. An option contract indexed to the spot market price – that is, a contract that specifies prices  $(0, m)$ , where  $m$  is the *realized* second-period spot market price – achieves the first-best outcome by ultimately allocating the resource to the user with the highest second-period value.

The first-best allocation  $z^{**}$  is that which maximizes the joint payoffs to the seller and buyer given realized demand level  $D$  and market price  $m$ .

**Proposition 3.1** *The first-best allocation is*

$$z^{**}(D, m) = \begin{cases} a(m; D) & \text{if } a(m; D) \leq K \\ K & \text{otherwise} \end{cases}$$

where  $a(x; D) = \{z \in \mathbb{R}_+ : u_z(z, D) = x\}$  and  $K$  is the total available supply of the good under contract.

---

<sup>5</sup>More generally, the option price is set equal to the seller's marginal opportunity cost of a sold option. For instance, under a model with extended production lead times, supply that is reserved under an option may need to be foregone as a production input thereby reducing potential second-period profits. In this case, there is a positive opportunity cost of a sold option. For simplicity, we have assumed here that any supply freed up from uncalled options can be directly offloaded onto the spot market, e.g., a very short production lead time. Hence the opportunity cost of a sold option – and the optimal option price – is zero.

*Proof.* The first-best allocation  $z^{**}$  is the solution to the following constrained maximization problem:

$$\begin{aligned} \max_{z \in [0, K]} \{u(z, D) + m(K - z)\} \\ \text{s.t. } z \leq K \end{aligned}$$

The Lagrangian is given as

$$L = u(z, D) + m(K - z) - \lambda(z - K)$$

Taking the FOC we have,  $u_z(z, D) - m - \lambda = 0$ , where  $\lambda = 0$  when the constraint  $z \leq K$  is non-binding by complementary slackness. The constraint is non-binding when  $a(m) \leq K$ , which implies  $u_z(z, D) = m$  and hence  $z^{**} = a(m; D)$  (under our concavity assumption on  $u$ ), where  $a(m; D) = \{z \in \mathbb{R}_+ : u_z(z, D) = m\}$ . Otherwise, the constraint is binding and hence  $z^{**} = K$ . ■

Proposition 3.1 establishes the optimality of marginal cost pricing in the second period. The marginal cost here is equal to the *realized* market price  $m$ , which is the cost (to the seller or, more broadly, society) of a single unit of consumption.

Now we suppose that option contracting is the mechanism through which a resource allocation  $z$  is to be determined, and we look for the price pair  $(p, s)$  that maximizes joint surplus under contracting. The assumption is that a single strike price  $s$  is set in the contract, e.g., no indexing of the strike price. Proposition 3.2 establishes the contract prices  $(p, s)$  and quantity  $Q$  that maximize the *expected* joint surplus. This is the ex-ante efficient contract.

**Proposition 3.2** *The ex-ante efficient option contract is given by  $(p, s) = (0, \bar{m})$ , where  $Q$  is determined by  $Q = [h(p + s(1 - F))]_+$  and  $h$  is the inverse function defined by  $h(x; s, D) = \{Q \in \mathbb{R}_+ : \bar{u}(Q; s, D) = x\}$  and  $\bar{u}(Q; s, D) = \int_{\bar{D}(s, Q)}^{\kappa} u_Q(Q, D) dF(D)$ .*

The ex-ante efficient contract sets an option price of zero and a strike price equal to the seller's expected opportunity cost on the spot market, which is the expected spot market value of the good  $\bar{m}$ . The contract does not achieve first-best, as the second-period allocation achieved under the rigid strike price of  $\bar{m}$  may differ from that achieved under the optimal allocation given in Proposition 3.1. Theorem 3.1 below summarizes this result, the proof of which relies on a first-order stochastic dominance argument.

**Theorem 3.1** *Any contract apart from the option contract with an execution price indexed to the realized spot market price  $m$ , i.e., any contract apart from that which sets prices  $(p, s) = (0, m)$ , is ex-post inefficient.*

*Proof.* From Proposition 3.1 we know that any contract that does not set the transaction price equal to the realized spot market price fails to maximize surplus in the second-period. For a contract with an execution price not equal to  $m$ , the second period payoff will therefore always be below that achievable at  $m$ . If the second period payoff in every state of the world is lower, then by FOSD we conclude that the *expected* payoff is lower. Hence a contract that names prices other than  $(0, m)$  is ex-ante inefficient. ■

REMARK 4 Regarding the welfare-maximizing contract, we observe that at a strike price of  $\bar{m}$ , the buyer will execute options in the second period up to the point that his marginal value of executing options  $u_z$  is equal to  $\bar{m}$ . In other words, the buyer's expected marginal value with respect to the number of options he exercises is constant:  $\bar{m}$ . The seller's expected marginal utility with respect to the number of supply units (options) he is left with is also constant:  $\bar{m}$ . The result in Proposition 3.2 is an incarnation of the familiar Borch rule (from Borch (1962)), which states that the optimal coinsurance contract is that in which the *ratio* of the marginal utilities of money are constant across all states of nature for the contracting parties. If one party's marginal utility is constant across states of nature, the Borch rule implies that the other party must then also have a constant marginal utility across all states of nature. Here, the seller's expected marginal utility (of one unit of supply) is constant and equal to the expected spot market price  $\bar{m}$ . Hence, under the optimal coinsurance contract the buyer's expected marginal utility is also constant. □

**Example 1.** In the following example we assume that the buyer minimizes a quadratic shortage cost function. The notation follows that introduced in the previous section. The first-best allocation is achieved according to the following optimization problem:

$$\begin{aligned} \max_{z \in [0, K]} \{ & -c(z - D)^2 + m(K - z) \} \\ \text{s.t. } & z \leq K \end{aligned}$$

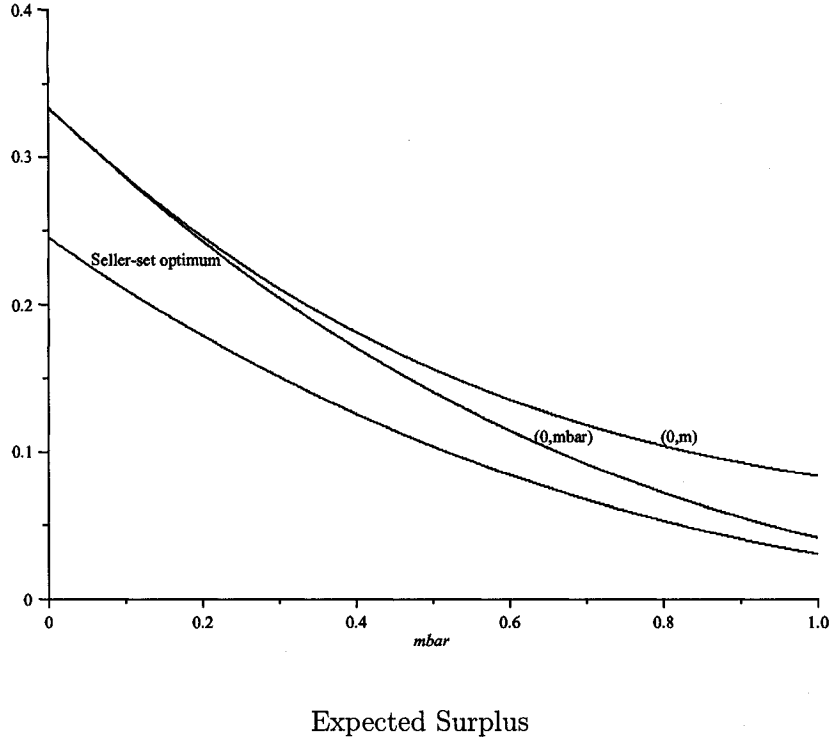


Figure 3.1: Expected Surplus

The FOC for the Lagrangian is given as

$$\frac{\partial L}{\partial z} = -2c(z - D) - m - \lambda = 0. \quad (3.3)$$

Solving (3.3) we find

$$z^{**} = \begin{cases} D - \frac{m}{2c} & \text{if } D \leq K + \frac{m}{2c} \\ K & \text{otherwise} \end{cases}$$

where

$$\lambda = \begin{cases} 0 & \text{if } D \leq K + \frac{m}{2c} \\ -2c(K - D) - m & \text{otherwise} \end{cases}$$

We note that if the upper bound of the support of demand  $\kappa$  is less than  $K + \frac{m}{2c}$  then the capacity constraint never binds and  $z^{**} = D - \frac{m}{2c}$ . Specifically, if  $\kappa > K + \frac{\beta}{2c}$ , where  $\beta$  is the upper bound of the spot market support, the constraint remains non-binding.

Given our expression for the optimal number of options that the buyer should be allocated  $z^{**}$  dependent on the realization of  $\tilde{D}$  and  $\tilde{m}$  (where  $K - z^{**}$  will be offloaded onto the market), we can ask, *what strike price would we have to set to get the buyer to execute the optimal number of options  $z^{**} = K - \frac{m}{2c}$ ?*

The buyer's second-period maximization problem is given as

$$\max_{z \in [0, K]} \{-c(z - D)^2 - sZ\}$$

Hence, taking the FOC (and letting  $z^{**}$  denote the buyer's optimal  $z$  given  $s$ ) we can then solve to find the  $s$  that implements  $z^{**}$ :

$$z = D - \frac{s}{2c} = D - \frac{m}{2c}$$

This implies  $s = m$ .

□

### 3.4 Conditions for Contracting

In general the two parties voluntarily enter into an option contract whenever gains from trade are positive. If gains from trade are positive, then there exists a price pair  $(p, s)$  such that both the seller and the buyer are better off under contracting. The general condition is provided in Lemma 3.1.

**Lemma 3.1** *An option contract of the form  $(p, s, Q)$  is entered into voluntarily if and only if*

$$U(Q, s; \tau) + sy - \bar{m}(K - y) \geq \Pi_0 + \bar{U}_0$$

where  $Q(p, s, K)$  and  $y(p, s, K)$  are the number of options held by the buyer and the number of options that the seller expects the buyer to exercise, respectively, and  $\Pi_0 = \bar{m}K$  is the seller's reserve value and  $\bar{U}_0 = \Phi_2$  is the buyer's reserve value. ■

#### 3.4.1 Contracting without an Outside Option

We first consider the case where the buyer has no outside option. Lemma 3.2 below establishes the necessary and sufficient condition for contracting – namely that the buyer's marginal value of a unit of supply in the second period is greater than the expected spot market price for at least one demand realization  $D$ . If the buyer is willing to pay at least  $\bar{m}$  to execute an option in some future state of the world, then the strike price can be set greater than  $\bar{m}$ . This in turn implies that the seller wants to contract. Succinctly, the necessary



and sufficient condition for contracting is simply that the buyer values a unit of supply more highly than the seller in some future state of the world. Contracting will then provide both the buyer and the seller with a higher expected value: the seller from the expectation that he can offload some supply at a price above the expected spot market price, and the buyer from the expectation that he can avoid losses with a higher marginal cost than option exercise.

**Lemma 3.2** *The marginal value of at least one option lying above  $\bar{m}$ , e.g.,  $u_Q(0, D) > \bar{m}$ , for some  $D \in [0, \kappa]$ , where  $\kappa$  is the upper bound of the support of  $D$ , is a necessary and sufficient condition for contracting (when the buyer has no outside option).*

We note that Lemma 3.2 does not predict the actual prices at which contracting will occur, nor does it preclude a positive contract price  $p^*$ .

### 3.4.2 Contracting with an Outside Option at a Fixed Price

The introduction of an outside option at a fixed price  $p_a$  imposes an additional restriction: the buyer will never be willing to enter into a contract with a strike price above the price of the outside option, e.g., we must have  $s < p_a$  (with  $p \geq 0$ ). The necessary and sufficient condition for contracting established in Lemma 3.2 has to be modified. The new conditions are as follows:  $\bar{m} \leq p_a$  and  $\bar{m} < u_Q(0, D)$ . The expectation of the spot market price  $\bar{m}$  must be below the price of the outside option, since the strike price must also be below the price of the outside option, e.g.,  $\bar{m} < s < p_a$ , and the buyer's marginal value of obtaining one unit  $Q$  must be greater than  $\bar{m}$ .

### 3.4.3 Contracting with Joint Market Access: Certain Demand

Now we consider the case where the buyer has an outside option, i.e., there is joint access to the spot market. We begin with the case of *certain downstream demand* as presented in Wu et al. (2003).<sup>6</sup> In this case, the buyer's expected payoff function is given as

$$U(p, s, Q) = -pQ + \int_s^\beta (m-s)Q dF(m) + \int_0^{u'(Q)} [u(D(m)) - mD(m)] dF(m) + \int_{u'(Q)}^\beta [u(Q) - mQ] dF(m). \quad (3.4)$$

where  $D = \max[D(m), Q]$  and for any  $m < u'(Q)$  it will be  $D = \max[D(m), Q] = D$ , since  $u^{-1}(u'(Q)) = Q$  and  $u'^{-1}$  is a decreasing function such that  $m < u'(Q)$ . This im-

---

<sup>6</sup>In Wu et al.'s notation, the option price is denoted as  $s$ , the strike price as  $g$ , the spot market price as  $P_s$ , demand as  $D$  and the quantity of options as  $Q$ . We retain our own notation here for consistency.

plies  $u'^{-1}(m) > Q$ . Also, we observe that in the case of certain demand, the buyer's execution strategy is a 'bang-bang' strategy: he either executes all  $Q$  options, if the spot market price realizes above the strike price  $s$ , and otherwise none (instead filling demand on the spot market). The buyer never holds more options  $Q$  than he would exercise at  $s$ .

Wu et al. introduce the *effective price function*, defined as follows for some price  $p$

$$G(p) = \int_0^p (1 - F(y))dy = \int_0^p mf(m)dm + p(1 - F(p)) = E[\min\{p, m\}]$$

where  $F(y)$  is the cdf of the spot market price  $\tilde{m}$  and the second equality is an application of integration by parts. The effective price function is used in deriving a final expression for the buyer's FOC from (3.4):

$$\begin{aligned} \frac{\delta U(p, s, Q)}{\delta Q} &= -p + \int_s^\beta (m - s)dF(m) + \int_{u'(Q)}^\beta (u'(Q) - m)dF(m) \\ &= -p + \int_0^\beta mdF(m) - \int_0^s mdF(m) - s(1 - F(s)) - \int_0^\beta mdF(m) + \int_0^{u'(Q)} mdF(m) + u'(Q)(1 - F(u'(Q))) \\ &= -p - G(s) + G(u'(Q)) = 0 \end{aligned}$$

The result is summarized in the following theorem from Wu et al. (2003):

**Theorem 3.2** (*Buyer's optimal contracting policy*) Assume  $\lim_{x \rightarrow 0^+} u'(x) < \infty$ . When  $p + G(s) > G(u'(0))$ ,  $Q^*(p, s) = 0$ . Otherwise,  $Q^*(p, s)$  is determined by any of the following identities:

- i.  $Q^*(p, s) = D(G^{-1}(p + G(s)))$
- ii.  $E\{\min\{m, u'(Q^*(p, s))\}\} = p + E\{\min\{m, s\}\}$
- iii.  $E\{m - u'(Q^*(p, s))\}^+ = p + E\{\min\{m, s\}\}$
- iv.  $u'(Q^*) = p + s + \int_s^{u'(Q)^*} F(m)dm$
- v.  $G(u'(Q^*)) - G(s) = p$

*Proof.* See Wu et al. (2003).

In the case of certain downstream demand, feasibility of contracting requires that the seller face some potential loss on the spot market. Wu et al. represent this potential loss as a constant factor that denotes the total percentage of supply that can be offloaded onto the

spot market, denoted here as  $r$ . The necessary (and sufficient) condition for contracting is  $r < 1$ , where the seller's expected payoff function is given as follows

$$\Pi = pQ + \int_s^\beta (s-b)QdF(m) + \int_0^s [r(m-b)_+K]dF(m) + \int_s^\beta [r(m-b)_+(K-Q)]dF(m) - \beta K. \quad (3.5)$$

The seller's expected payoff function can be rewritten in terms of the effective price function:

$$\begin{aligned} V &= pQ + (s-b)(1-F(s))Q + rK \left( \int_0^\beta mdF(m) - \int_0^s mdF(m) - \int_b^\beta bdF(m) \right) \\ &\quad - \chi(s-b) \left( \int_0^\beta mdF(m) - \int_0^s mdF(m) - \int_s^\beta bdF(m) \right) Q \\ &\quad - \chi(b-s) \left( \int_0^\beta mdF(m) - \int_0^b mdF(m) - \int_b^\beta bdF(m) \right) Q - \beta K \\ &= pQ + rK[\bar{m} - G(b)] - \chi(s-b)[\bar{m} - G(s)] - \chi(b-s)[\bar{m} - G(b)] - \beta K \\ &= pQ + (s-b)(1-F(s))Q - r\chi(s-b)(s-b)(1-F(s))Q - [\bar{m} - G(\max(s,b))]Q + rK[\bar{m} - G(b)] - \beta K. \end{aligned} \quad (3.6)$$

The seller's expected opportunity cost on the spot market from (3.6) is  $rQ(\bar{m} - G(b))$ . Theorem 3.3 establishes the necessary and sufficient condition for contracting that  $r < 1$ .

**Theorem 3.3** (From Wu et al.) Given  $Q$  as determined by (19). Then  $Q > 0$  implies

$$0 \leq r < \frac{G(U'(0)) - G(b)}{\bar{m} - G(b)} < 1.$$

*Proof.* The seller's revenue from the contract market must offset her opportunity cost on the spot market, i.e.,  $rQ(\bar{m} - G(b))$ . This gives us

$$r \leq \frac{p^*}{\bar{m} - G(b)} = \frac{G(u'(Q)) - G(b)}{\bar{m} - G(b)} \leq \frac{G(u'(0)) - G(b)}{\bar{m} - G(b)} < 1.$$

The above equality holds due to Theorem 3.2. The second inequality holds due to the decreasing of  $u'$  and the increasing of  $G$  ( $G'(x) = 1 - F(x) > 0$ ). The last inequality holds by the definition of the effective price function. Hence the proof.  $\blacksquare$

As we now demonstrate, the presence of uncertain downstream demand alters the conditions for feasible contracting, relaxing the dependence on the supplier's risk factor.

### 3.4.4 Contracting with Joint Market Access: Uncertain Demand

Now we consider the buyer's and seller's payoffs in the case of *uncertain* downstream demand. Under uncertain demand we no longer have a 'bang-bang' decision outcome where the buyer exercises  $Q$  or 0. Instead, for given realizations of demand the buyer will elect to exercise some fraction of the total options that he holds  $Q$ . The seller's payoff function is identical in form to that with certain demand, where we replace  $Q$  with  $y$ , where  $y = Q(1 - F) + \bar{g}$  is the expected number of options that the buyer will actually exercise:

$$V = pQ + (1 - F(s))(s - b)y + [\bar{m} - G(\max(s, b))]y + K[\bar{m} - G(b)] - \beta K$$

The seller's opportunity cost at the spot market is now given as  $y(\bar{m} - G(b))$  where we note that we have granted the seller full access to the spot market (there is no 'risk factor'  $r$ ).

The buyer's expected payoff function when facing uncertain downstream demand is given as

$$\begin{aligned} U(p, s, Q) = & -pQ + \int_s^\beta \left[ \int_0^{D(s)} u(0, D) dF(D) + \int_{D(s)}^{\bar{D}(s, Q)} [u(g(s, D), D) - sg(s, D)] dF(D) \right] dF(m) + \\ & \int_s^\beta \int_{\bar{D}(s, Q)}^{\bar{D}(m, Q)} [u(Q, D) - sQ] dF(m) dF(D) + \int_s^\beta \int_{\bar{D}(m, Q)}^\kappa [u(g(m, D), D) \\ & - sQ - m(g(m, D) - Q)] dF(m) dF(D) \end{aligned}$$

A number of the integral terms cancel when we take the derivative with respect to  $Q$  and we are left with

$$\frac{\delta U}{\delta Q} = -p + \int_s^\beta \int_{\bar{D}(s, Q)}^{\bar{D}(m, Q)} [u_Q(Q, D) - s] dF(m) dF(D) + \int_s^\beta \int_{\bar{D}(m, Q)}^\kappa [-s + m] dF(m) dF(D)$$

where the last term can be rewritten in terms of Wu et al's effective price function as follows

$$\begin{aligned} & \int_s^\beta \int_{\bar{D}(m, Q)}^\kappa [-s + m] dF(m) dF(D) \\ = & \int_0^\beta m(1 - F(\bar{D}(m, Q))) dF(m) - \int_0^s m(1 - F(\bar{D}(m, Q))) dF(m) - \int_s^\beta s(1 - F(\bar{D}(m, Q))) dF(m) \end{aligned}$$

Hence, the buyer's FOC is given as

$$p + (1 - F)(\bar{m} - G(s)) + \int_s^\beta \int_{\bar{D}(s, Q)}^{\bar{D}(m, Q)} [u_Q(Q, D) - s] dF(m) dF(D) = 0. \quad (3.7)$$

In Theorem 3.4 below we establish that under uncertain demand we can find conditions under which contracting will occur *even when the seller has unmitigated access to the spot market*. As in the case of certain demand, we begin by considering the seller's revenue requirement. The seller's revenue from contracting must offset her opportunity cost on the spot market:  $py \geq y(\bar{m} - G(b))$ .

**Theorem 3.4** *For large enough  $u_Q(Q, D)$  (where  $u_Q(Q, D)$  can be arbitrarily large) we obtain feasible contracting between the seller and the buyer even when the seller has unmitigated access to the spot market.*

*Proof.* For contracting, we need that the seller's opportunity cost on the spot market is offset by the revenue from the contract market:  $py \geq y(\bar{m} - G(b))$ . This implies the following relation:  $1 \leq \frac{p^*}{\bar{m} - G(b)} = \frac{(1-F)(\bar{m} - G(s)) + \int_s^\beta \int_{D(s, Q)}^{D(m, Q)} [u_Q(Q, D) - s] dF(m) dF(D)}{\bar{m} - G(b)}$ , where the equality holds by (3.7). We can satisfy this condition by, for example, fixing  $s$  and selecting some arbitrarily large  $u_Q$ . ■

In Section 3.8, we see that when the buyer faces a large enough shortage cost, e.g., a large enough weight  $c$  in the quadratic shortage cost function  $-cu(z, D)$ , we do indeed get contracting (even when the market friction  $\tau$  is eliminated). Increasing  $c$  corresponds directly to an increase in  $u_Q(Q, D)$  in the theorem above.

### 3.5 Contracting without an Outside Option

Without an outside supply option, the buyer relies entirely on the seller for supply to meet his downstream demand. This scenario applies rather generally to cases where an intermediary in a given market looks to expand supply and faces a single upstream supplier. The intermediary may hold existing supply contracts with other suppliers and may, in fact, be contracting to meet excess demand, i.e., demand beyond that which he can serve with existing supply. In the water market, where expansion of supply often entails construction of additional infrastructure, an intermediary in the contract market may have very limited outside supply options in the short-term.

A contract  $(p, s, Q)$  is set when the seller, acting as a Stackelberg leader, offers a price pair  $(p, s)$ , and the buyer responds with demand for a quantity of options  $Q(p, s)$ . In a complete information setting, the number of options demanded at a given price pair  $(p, s)$  is known to the seller in advance. More precisely, the buyer's demand  $Q(p, s)$  is an input to the

seller's profit maximization problem, where  $Q(p, s)$  is the solution to the buyer's expected utility maximization problem. The structure of the optimal contract is investigated below beginning with the buyer's optimization problem and the characterization of the optimal  $Q$ . In the second period, the buyer observes the realized demand  $D$ . Facing an exercise price  $s$  and an available quantity of options  $Q$ , the buyer solves the following problem:

$$\max_{z \leq Q} \{u(z, D) - sz\} \quad (3.8)$$

where  $u(z, D)$  is the buyer's gross payoff when delivering the amount  $z$  given a demand of  $D$ . For instance, the buyer's payoff may be a function of his shortfall in meeting demand, i.e., a shortage cost. In the example developed in Section 3.8, the buyer's utility function is the quadratic shortage cost function  $u(z, D) = -c(z - D)^2$  (for  $z < D$  and zero otherwise), where

$$z(Q, s, D) = \begin{cases} \max(0, g(s, D)), & \text{if } g(s, D) \leq Q, \\ Q, & \text{otherwise,} \end{cases}$$

and  $g(s, D) = \{z : u_z(z, D) = s\}$ .

The buyer's expected utility of contracting for  $Q$  options in the first period, obtained by substituting  $z(Q, s, D)$  into the objective function in (3.8) and taking the expectation with respect to  $D$ , is therefore

$$U(Q, s) = E[u(z(Q, s, D), D) - sz(Q, s, D)].$$

We observe that  $U(Q, s)$  is generally increasing.

At a price pair  $(p, s)$ , the buyer's first-period problem is to maximize his expected payoff

$$\max_{Q \geq 0} \{U(Q, s) - pQ\}.$$

The realization of demand will result in one of three actions by the buyer: (1) demand is low, and the buyer will elect not to exercise any options, (2) demand is in a 'mid-region,' and the buyer will execute some fraction of the options, and (3) demand is high, and the buyer will execute all  $Q$  options. Our expression for the Buyer's expected utility in Period 1 reflects these three cases:

$$\begin{aligned} U(Q, s) - pQ &= \int_0^{D(s)} u(0, D) dF(D) + \int_{D(s)}^{\bar{D}(s, Q)} [u(g(s, D), D) - sg(s, D)] dF(D) \\ &\quad + \int_{\bar{D}(s, Q)}^{\infty} [u(Q, D) - sQ] dF(D) - pQ \end{aligned}$$

where  $\underline{D}(s) = \{D : g(s, D) = 0\}$  and  $\bar{D}(s, Q) = \{D : g(s, D) = Q\}$ .

Taking the buyer's *FOC* yields an expression for the optimal  $Q$  in terms of the inverse function  $h$ :

$$Q^*(p, s) = [h(p + s(1 - F(\bar{D})))]_+$$

where  $h(x; s, D) = \{Q \in \mathbb{R}_+ : \bar{u}(Q; s, D) = x\}$  and  $\bar{u}(Q; s, D) = \int_{\bar{D}(s, Q)}^{\kappa} u_Q(Q, D) dF(D)$ . This leads us to the two lemmas below. Lemma 3.3 characterizes the buyer's demand as decreasing in both prices, a direct result of the decreasingness of  $h$  in  $p$  and  $s$ . Lemma 3.4 defines the ratio of change in  $Q$  with respect to  $p$  and  $s$ .

**Lemma 3.3** *The quantity of options purchased by the seller  $Q$  is decreasing in both  $p$  and  $s$ .*

**Lemma 3.4** *The ratio governing the change in  $Q$  relative to  $p$  and  $s$  respectively is inversely proportional to the probability that demand  $\bar{D}$  will realize such that **more than**  $Q$  options are desired and is given as  $\frac{Q_p}{Q_s} = \frac{1}{1 - F(\bar{D})}$ .*

Lemma 3.4 provides some insight into the relative trade-off between the upfront fee  $p$  and the strike price  $s$  from the buyer's perspective. Demand  $Q$  is more sensitive to an increase in the upfront fee than an increase in the strike price  $s$ . What is also evident from Lemma 3.4 is the existence of multiple price pairs  $(p, s)$  at which the buyer demands the same  $Q$ . Although the buyer is indifferent between points on his iso-demand curve, there is a unique optimum for the contracting problem, dictated by the seller.

We now turn to the seller's problem:<sup>7</sup>

$$\max_{p, s} \{pQ(p, s) + (s - b)y + E[(\tilde{m} - b)(K - y)]\}.$$

where  $y(Q, s) = E[z(Q, s, D)]$  is the number of options that the seller expects the buyer to exercise, given that the buyer purchases  $Q$  (where clearly  $y < Q$ ). The expression for  $y$  is given as

$$y(Q, s) = \int_{\underline{D}(s)}^{\bar{D}(s, Q)} g(s, D) dF(D) + \int_{\bar{D}(s, Q)}^{\kappa} Q dF(D) \quad (3.9)$$

<sup>7</sup>If  $Q = K$ , then the seller's problem is simpler:

$$\max_s \{(s - b)y(K, s) + E(\tilde{m} - b)(K - y(K, s))\},$$

which yields the optimal exercise price  $s = s(K, m)$ . The optimal option price is then  $p = p(K, m)$ , given as the solution of  $K = h(p, s(K, m))$ .

The seller's FOCs characterize the optimal  $p$  and  $s$  when the payoff function is concave. The seller's FOCs:

$$V_p = pQ_p + Q + (s - m)y_p = 0 \quad (3.10)$$

$$V_s = pQ_s + y + (s - m)y_s = 0, \quad (3.11)$$

where  $y_p$  and  $y_s$  are given as

$$y_p = Q_p(1 - F(\bar{D})) \quad (3.12)$$

$$y_s = Q_s(1 - F(\bar{D})) + \int_{D(s)}^{\bar{D}(s,Q)} g_s(s, D) dF(D).^8 \quad (3.13)$$

### 3.5.1 Seller Optimum

We introduce two lemmas characterizing the optimal prices and two assumptions that are needed for the proof of Theorem 3.5 below, establishing the prices in the seller-optimal contract. (See Appendix for proofs.)

**Lemma 3.5** *The seller charges  $p \geq 0$  and  $s > 0$ .*

**Lemma 3.6** *The optimal prices  $p$  and  $s$  are bounded above, assuming  $V'(0, D) < \infty$ .*

We introduce two assumptions needed in the proof of Theorem 3.5 below:

$$\text{A1. } V_{pp} + V_{ss} \leq 0$$

$$\text{A2. } V_{pp}V_{ss} - V_{ps}^2 \leq 0$$

Assumption A1 and A2 ensure joint concavity of the seller's profit function  $V$ , where  $H(V) = \frac{\partial^2 V(p,s)}{\partial(p,s)^2}$  is the Hessian of the profit function with  $\det H(V) = V_{pp} + V_{ss}$  and  $\text{trace } H(V) = V_{pp}V_{ss} - V_{ps}^2$ , so that under A1 and A2 both eigenvalues of  $H(V)$  are negative.

**Theorem 3.5** *(Optimal option price and strike price) Under assumptions A1 and A2, the seller optimally sets prices as follows:*

$$s^* = \bar{m} - \frac{\bar{g}}{\int_D^{\bar{D}} g_s dF(D)}$$

---

<sup>8</sup>We note that if  $g_s(s, D) = g_s(s)$ , i.e.,  $g_s$  independent of  $D$ , then the above expression simplifies to  $y_s = Q_s(1 - F(\bar{D})) + g_s(F(\bar{D}) - F(\underline{D}))$ .



and

$$p^* = \frac{-(s - \bar{m})(1 - F)}{1 - 1/\varepsilon_p}.$$

where  $\bar{g} = \int_{D(s)}^{\bar{D}(s,Q)} g(s, Q) dF(D)$ , we have written  $(1 - F)$  in place of  $(1 - F(\bar{D}(s, Q)))$ , and  $\varepsilon = Q_p \frac{p}{Q}$ .

REMARK 5 Assumptions A1 and A2 are quite general assumptions. Assumption A1, for instance, is satisfied whenever the profit function  $V$  is concave in both  $p$  and  $s$  respectively. In an attempt to provide more detailed necessary conditions for joint concavity, we might examine  $V_{pp}$  and  $V_{ss}$  directly. We have  $V_{pp} = [2 - f(\bar{D})\bar{D}_Q Q_p]Q_p + [p + (1 - F)(s - m)]Q_{pp}$ , where the first term is negative under the assumption that  $f(\bar{D}) > 0$ , since then  $[2 - f(\bar{D})\bar{D}_Q Q_p] > 0$  and  $Q_p < 0$ ). The sign of the second term depends on the sign of  $Q_{pp}$ :  $Q_{pp} \leq 0$  is a sufficient, but not necessary, condition for  $V_{pp} \leq 0$ . However, such an assumption (that  $Q_{pp} < 0$ ) is unsatisfactory. As we will see, even for a simple example we find  $Q_{pp} < 0$  does not hold, *yet* joint concavity of  $V$  in  $(p, s)$  does hold (and, hence, the rules for setting  $p^*, s^*$  established in Theorem 3.5 apply and can be used to characterize the global maximum). Throughout this chapter we elect to derive optimal pricing rules *under the assumption that joint concavity holds*, an assumption that can be checked directly in practice. Finally we note that when joint concavity *does not hold*, a solution to the optimal contracting problem is still readily obtained using numerical methods, as a consequence of compactness and the continuity of the buyer's and seller's payoff functions (which yield a well-defined global optimization problem).  $\square$

In the conventional single-product case, a monopolist prices in the *inelastic* part of the demand curve, i.e., where the price elasticity of demand is greater than unity. Letting  $\varepsilon_p$  denote the price elasticity of demand, we have, in the standard setting, the inverse elasticity rule:

$$1 \geq \frac{p - MC(D(p))}{p} = \frac{1}{\varepsilon_p}, \quad (3.14)$$

where  $MC(D(p))$  is the monopolist's marginal cost function. Solving for  $p$  in (3.14), we have

$$p = \frac{MC}{1 - 1/\varepsilon_p}. \quad (3.15)$$

Contrary to the standard rule, we see that in the contracting problem the seller (who is a monopolist in our bilateral setting), prices the option in the *elastic* part of the buyer's demand curve. Corollary 1 below states this result.

**Corollary 1** *The seller optimally prices in the elastic part of the demand curve. That is, the option price elasticity of demand  $\varepsilon_p = Q_p \frac{p}{Q} \leq 1$ .*

The explanation for this behavior comes from the monopolist's role as a multiproduct – in this case two-product – monopolist. As such, he is interested in lowering the price of the first good (here, the option) to encourage (possible) consumption of the second good (the strike). The pricing rule formalized in Theorem 3.5 can be interpreted in light of the standard theory of multiproduct pricing. In particular, our option pricing problem can be recast as a multiproduct pricing decision where the “products” are the option (in the first period) and the strike in the second period. From Tirole (1988, p. 70), the monopolist's first order conditions for the multiproduct pricing decision yield the following expression

$$\frac{p_i - MC_i}{p_i} = \frac{1}{\varepsilon_{ii}} - \sum_{j \neq i} \frac{(p_j - MC_j) D_j \varepsilon_{ij}}{p_i D_i \varepsilon_{ii}} \quad (3.16)$$

where  $p_i$  and  $MC_i$  are the price and marginal cost of product  $i$ , respectively,  $\varepsilon_{ii}$  is the own price elasticity of demand for product  $i$ ,  $D_j$  is the demand for product  $j$ , and  $\varepsilon_{ij}$  is the cross-elasticity of demand for product  $j$  with respect to the price of product  $i$ .

The Lerner index (the LHS of (3.16)) is equal to the inverse of the own price elasticity of demand *minus* the term  $\sum_{j \neq i} \frac{(p_j - MC_j) D_j \varepsilon_{ij}}{p_i D_i \varepsilon_{ii}}$ . Tirole (1988) observes that, in the case where the products are substitutes, this term will be negative:  $\frac{\partial D_j}{\partial p_i} > 0$  and, hence,  $\varepsilon_{ij} < 0$ . The Lerner index then exceeds the inverse of the own elasticity of demand. In the case of complements, where  $\frac{\partial D_j}{\partial p_i} < 0$ , the additional term on the RHS is positive, and the Lerner index is less than the own price elasticity of demand.<sup>9</sup>

We now show that the option price  $p$  is exactly that predicted in the monopolist multiproduct pricing problem. The key observation here is that in the option pricing problem, letting subscript one denote the option and subscript two denote the strike, the buyer holds exactly as many strikes as the number of options purchased. Hence,  $\frac{\partial D_2}{\partial p_2} = \frac{\partial D_1}{\partial p_1}$ , where the buyer “demands” as many strikes as options purchased.

Using this observation, algebraic manipulation of (3.16) above yields following:

$$p_1 = \frac{-(p_2 - MC_2)}{1 - 1/\varepsilon_{11}} \quad (3.17)$$

which is exactly the option price  $p$ , with  $p_2 = s$ ,  $MC_2 = m$ , the marginal opportunity cost of selling an option, and  $\varepsilon_{11} = \varepsilon_p$ . (Recall that the single-product monopoly price

<sup>9</sup>Andriychenko et al. (2006) point out the fact that whether the price charged by the monopolist for product  $i$  is indeed higher, or lower, than that charged by the duopolist is ambiguous. Although the additional term on the RHS of (3.16) is positive in the case of complementary goods and, hence, the inverse own elasticity of demand is greater than the Lerner index, the expressions for the monopolist and duopolist don't admit direct comparison due to differences in the optimal elasticities, which in turn depend on the optimal prices. Andriychenko et al. demonstrate, for instance, that it is possible for the monopolist to price one of two complementary goods higher than the duopolist.

in (3.15) simply had  $\frac{C_1'}{1-1/\varepsilon_1}$  on the RHS. The two-product monopolist sets the price of good one dependent on the price and marginal cost of good two. If he optimally prices good two above its marginal cost, then he will price good one in the inelastic part of the demand curve ( $\varepsilon_{11} \leq 1$ ). Whereas if he prices good 2 below marginal cost, then he will price good one in the elastic part of the demand curve (as a single-product monopolist would).

As we see later, this pricing rule does not hold once the buyer gains access to the spot market. Under conditions which are less favorable to the seller, the seller exploits his monopolistic power more strongly and reverts to pricing in the inelastic part of the demand curve.

**REMARK 6** In the case where demand is certain, we get exactly the monopolist pricing rule. There is no option contract, per se, but a forward contract for delivery in the second period with a price of  $\hat{p} = \frac{m}{1-1/\varepsilon_p}$ , obtained by solving the seller's problem for the case where the buyer optimally orders  $Q$  to satisfy  $u'(Q, D) = s + p$  (assuming  $u(Q, D) > u(0, D)$ ). That is, knowing demand  $D$  and unaffected by the uncertain spot market price (to which he does not have access), the buyer optimally sets his marginal value equal to his marginal cost. The seller's maximization problem is

$$\max_{p,s} \{pQ(p, s) + (s - b)y + E(\tilde{m} - b)_+(K - y)\}.$$

In the case of certain demand and no spot market access,  $y = Q$  and the buyer (and seller) weigh  $s$  and  $p$  the same.<sup>10</sup> □

### 3.5.2 Social Optimum

From our discussion in Section 3.3, we know that this contract does not maximize joint surplus. In comparison to the socially optimal option contract, the seller-set contract charges an upfront fee and inflates the strike price. From Lemma 3.5, we know that  $p > 0$ , which implies that the number of options held by the buyer  $Q(p^*, s^*)$  will be below the socially-optimal level. Figure 3.5.2 illustrates this fact: the first-best allocation  $Q^{FB}$  is higher due to both the elimination of the upfront fee and the reduction in  $s$ .

To see that the number of options held under the seller-set contract is lower than the number held under the social-welfare-optimal contract, we consider the ex-ante welfare maximizing option contract that sets prices  $(0, \bar{m})$ . At these prices the buyer's optimal  $Q$  is given as the

<sup>10</sup>This would need to be adjusted if one introduced a discount rate for the second-period.

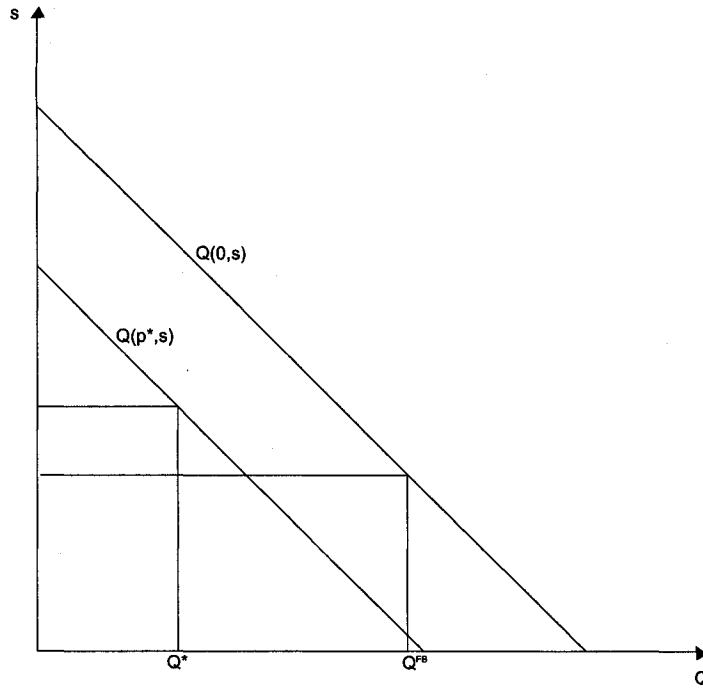


Figure 3.2: Comparison to the First-Best Allocation

solution to the following maximization problem:

$$\begin{aligned} \max_Q \int_0^{D(s)} u(0, D) dF(D) + \int_{D(s)}^{\bar{D}(s, Q)} u(g(D), D) dF(D) + \int_{\bar{D}(s, Q)}^{\kappa} u(Q, D) dF(D) \\ - \bar{m} \left[ \int_{D(s)}^{\bar{D}(s, Q)} g(D) dF(D) + \int_{\bar{D}(s, Q)}^{\kappa} Q dF(D) \right] + \bar{m}K \end{aligned}$$

from which we obtain the first order condition

$$\int_{\bar{D}}^{\kappa} u_Q(Q, D) = -\bar{m}(1 - F)$$

This implies

$$Q_{SW}^* = [h(\bar{m}(1 - F(\bar{D})))]_+$$

where  $Q_{SW}$  denotes the  $Q$  under the social welfare maximizing contract. That  $Q_{SW}^* = h(\bar{m}(1 - F(\bar{D}))) > h(p^* + s^*(1 - F(\bar{D}))) = Q^*$  is immediate, since  $s^* > \bar{m}$  by Theorem 3.5 and  $h$  decreasing.

Although the seller-set contract differs from the social-welfare maximizing contract (resulting in an additional loss of surplus), it is important to note that we can achieve the welfare-maximizing outcome through the introduction of transfers. The introduction of a non-linear

pricing scheme – namely, two-part tariffs – would facilitate social welfare maximization. With a two part-tariff, the seller maximizes total surplus and then confiscates the buyer's surplus through a fixed fee, or transfer payment.

The discussion in Section 3.8 below, in which we present a detailed example of the option pricing problem, further highlights the differences between the socially optimal contract and that struck between a profit-maximizing seller and a single buyer.

### 3.5.3 Example: Contracting without an Outside Option

The buyer is an intermediary supplying downstream demand and, as we assumed earlier, faces a quadratic shortage cost for all unmet demand:

$$u(z, D) = \begin{cases} -c(z - D)^2 & \text{for } z < D \\ 0 & \text{otherwise.} \end{cases}$$

The buyer's payoff function in the second-period is  $u(z, D) - sz$ , where  $z(Q, s, D) = \min\{[g(s, D)]_+, Q\}$  and  $g(s, D) = \{z : u_z(z, D) = s\}$ . For our example,  $u_z = -2(z - D)$  and, hence,  $g(s, D) = D - s/2$ .

For demand  $\tilde{D}$  uniformly distributed on  $[0, 1]$ , the optimal prices are given as

$$p^* = \frac{1}{25}\bar{m}^2 - \frac{4}{25}\bar{m} + \frac{4}{25} \quad (3.18)$$

$$s^* = \frac{4}{5}\bar{m} + \frac{2}{5}, \quad (3.19)$$

where  $\bar{m}$  is the mean of the spot market price  $\tilde{m}$ . Then the buyer's optimal contract order is

$$Q^*(p, s) = -\frac{1}{2}s + 1 - \sqrt{p}, \quad (3.20)$$

which is decreasing in both prices.

Figures 3.3 and 3.4 illustrate the following: (1) the option price  $p$  is decreasing in the expected spot market price  $\bar{m}$ , while the strike price  $s$  is increasing in  $\bar{m}$ , (2) the number of options purchased by the buyer  $Q$  is decreasing in the seller's reserve value, i.e., decreasing in  $\bar{m}$ , (3) the seller's payoffs are increasing in  $\bar{m}$ , while the buyer's payoffs are decreasing, and (4) the seller's surplus is always above the buyer's surplus, and both surpluses are decreasing in  $\bar{m}$ .

As proved in Theorem 3.5, the seller always charges a strike price  $s$  that is above the expected spot market price  $\bar{m}$ . As the relative strength of the seller's reserve (measured in terms of the expected spot market price per unit,  $\bar{m}$ ) increases, we see that the seller demands a

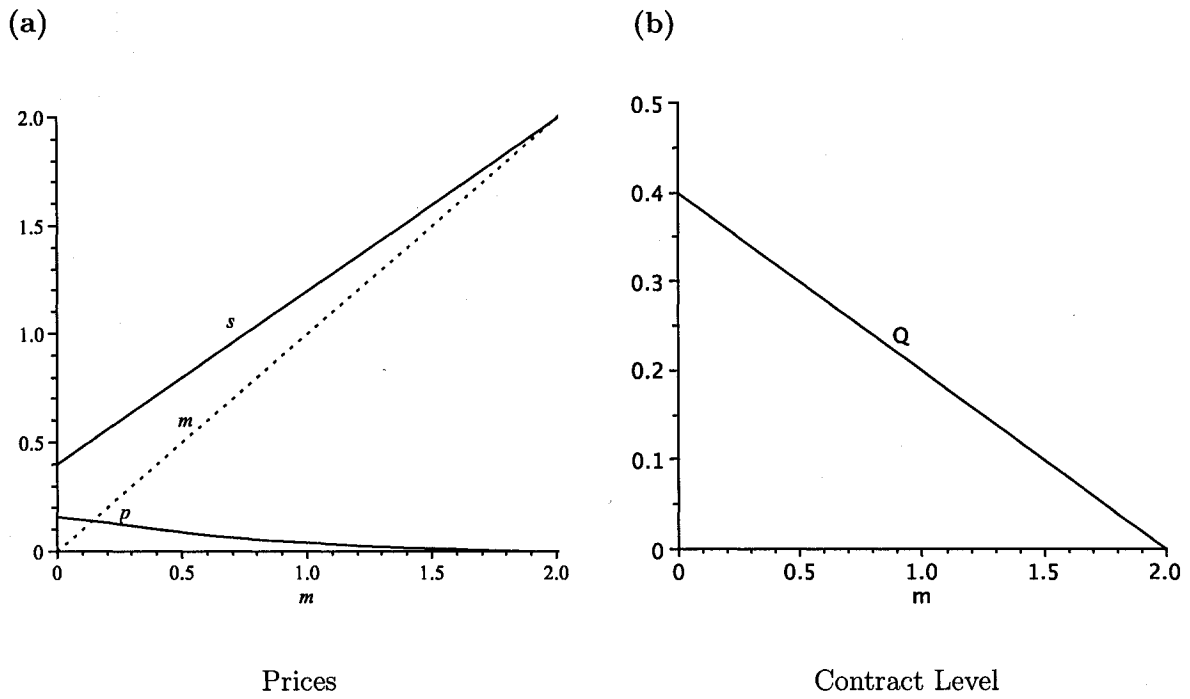


Figure 3.3: Sensitivity of the Contract Level and Prices to Expected Market Price

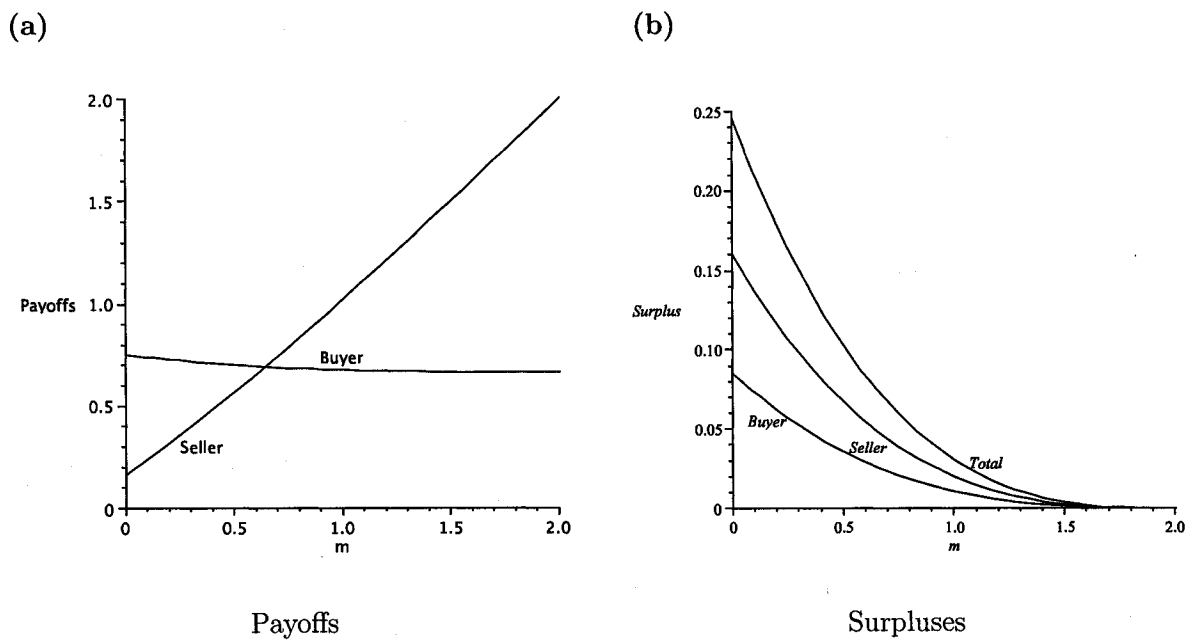


Figure 3.4: Sensitivity of Buyer's and Seller's Payoffs to Expected Market Price

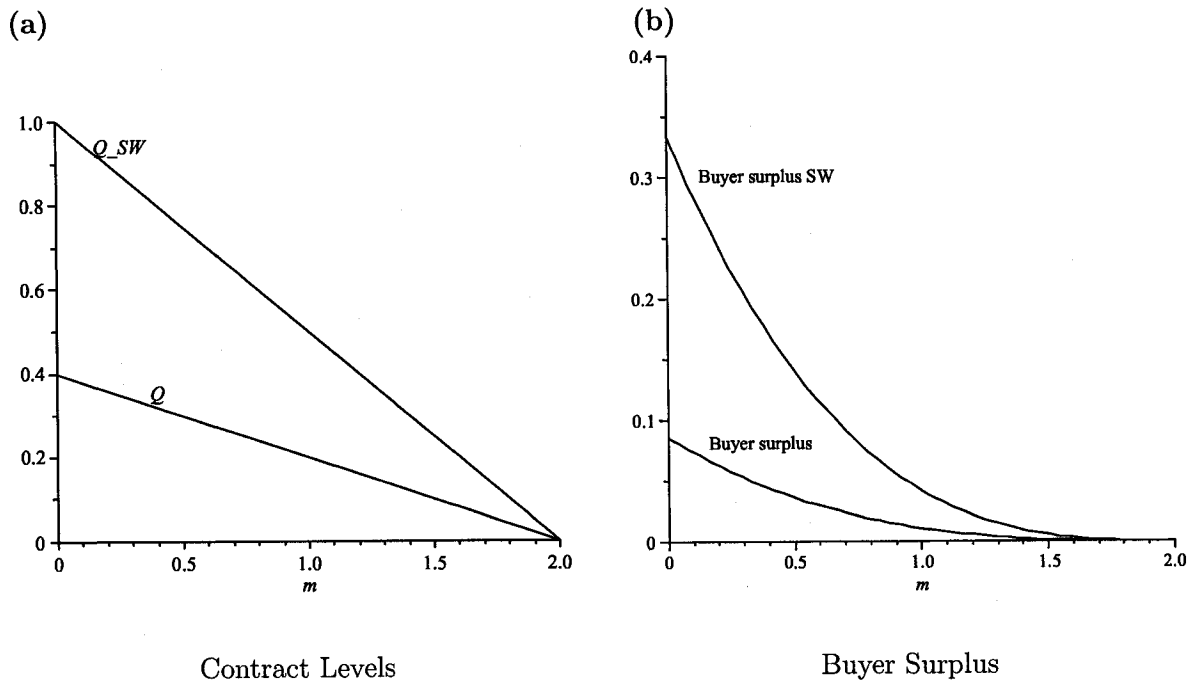


Figure 3.5: Comparison of Contract-Level and Payoff Under Profit Maximization and Social Welfare Maximization

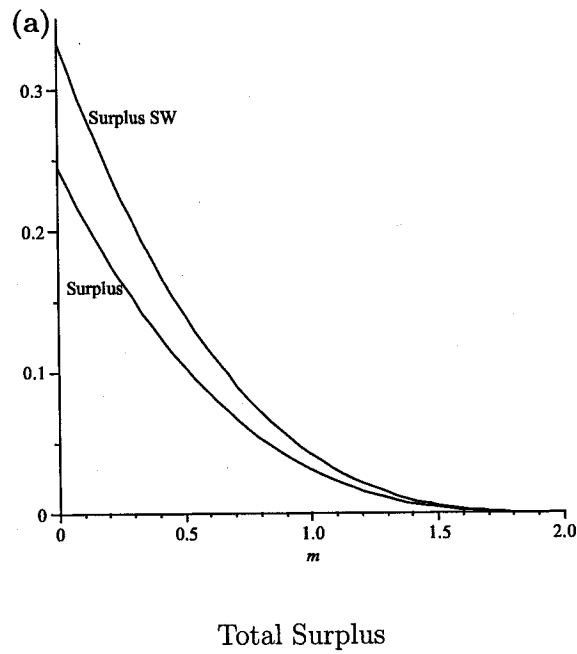


Figure 3.6: Deadweight Losses Under the Seller-Set Optimum

higher strike price, while decreasing the option price  $p$  (so as to ensure that the buyer still purchases a number of options, which have an expected payoff above the market price due to  $s > \bar{m}$ ). That  $p^*$  and  $s^*$  move in different directions is a result of the supermodularity: the seller's payoff function is supermodular in  $(p, -s)$  (equivalently in  $(-p, s)$ ).

The seller's and buyer's individual rationality constraints dictate that each will only enter into a contract if there is a net benefit to doing so. The seller's expected reserve payoff is equal to  $\bar{m}K$ , where  $K$  is the seller's capacity. Hence, for any  $p \geq 0$  and  $s \geq m$ , the seller's expected payoff is above his reserve. We can solve for the buyer's reserve payoff:  $\int_0^1 -(-d)^2 f(d) dD = -\frac{1}{3}$ . For any contract with an expected payoff greater than  $-1/3$ , the buyer is willing to contract. We see in Figure 3.4 that the buyer's individual rationality constraint is indeed satisfied.

Figures 3.5 and 3.6 illustrate the differences between the socially optimal contract and that obtained between a profit-maximizing seller and a single buyer. The optimal quantity of options available to the buyer at the social optimum is higher than that under the seller-set optimum. More precisely,  $Q_{SW}$  is exactly equal to the maximum number of options that the buyer would ever call at price  $\bar{m}$ : the buyer can meet all downstream demand *up to the point* where the marginal (opportunity) cost  $\bar{m}$  of doing so exceeds the marginal expected benefit. The buyer's surplus is higher at the social optimum, while the seller's surplus is lower. In fact, the seller's surplus is fully appropriated.

### 3.6 Contracting with an Outside Option

The buyer may have an alternative supply option available at a fixed price. This may be in the form of an alternative technology, i.e., if the buyer does not contract for supply he may go into production himself. In the water market this may be the option to rely on desalination (a short-term option *only if* the desalination plant is already operational). Alternatively, the outside option at a fixed price may represent an existing (flexible) supply contract. In the case where the intermediary faces a shortage cost for not meeting demand, the alternative supply option may be the cost of demand-side management programs such as water conservation programs. The presence of this supply option impacts the structure of the optimal contract *irrespective of whether the buyer actually expects to rely on it once the contract is signed*.

The effect of the outside supply option is essentially to tighten the buyer's individual ra-



tionality (IR) constraint. The buyer's IR constraint dictates that the buyer's utility from contracting (at the price pair  $(p, s)$  offered by the seller) must be higher than his utility from not contracting, or reserve value. If this constraint is not satisfied, we say the contracting problem is infeasible – a rational buyer will elect not to contract.

We revisit the buyer's problem, where the availability of a secondary supply option at a fixed price creates the possibility that the buyer can meet any downstream demand that he doesn't fill through the contract at a known price  $p_a$  in the second period:

$$\begin{aligned} U(Q, s, p_a) - pQ &= \int_0^{D(s)} u(0, D) dF(D) + \int_{D(s)}^{\bar{D}(s, Q)} [u(g(s, D), D) - sg(s, D)] dF(D) \\ &\quad + \int_{\bar{D}(s, Q)}^{\min\{\kappa, \bar{D}(p_a, Q)\}} [u(Q, D) - sQ] dF(D) + \\ &\quad \int_{\min\{\kappa, \bar{D}(p_a, Q)\}}^{\kappa} u(g(p_a, D), D) - sQ - p_a(g(p_a, D) - Q) dF(D) - pQ \end{aligned} \quad (3.21)$$

where we recall that the support of the buyer's demand distribution is  $[0, \kappa]$  and where  $\bar{D}(s) = \{D : g(s, D) = 0\}$ , and  $\bar{D}(s, Q) = \{D : g(s, D) = Q\}$ . (And, hence,  $\bar{D}(p_a, Q) = \{D : g(p_a, D) = Q\}$ .)

In the event that the buyer does not expect to rely on the outside option for any additional supply, i.e.,  $\min\{\kappa, \bar{D}(p_a, Q)\} = \kappa$ , which would generally be the case for a large enough  $p_a$  or a large enough  $Q$ , the expression for the buyer's expected utility collapses to that examined above for the case *without an outside option*. However, the presence of the outside option will still enter into the seller's problem through the buyer's individual rationality constraint.

The buyer's FOC yields the following characterization of the optimal  $Q(p, s)$ :

$$Q^* = [\hat{h}(p + s(1 - F(\bar{D}(s, Q)))) - p_a(1 - F(\bar{D}(p_a, Q)))]_+, \quad (3.22)$$

where  $\hat{h}(x) = \{Q \in \mathbb{R}_+ : \hat{u}(Q; s, D) = x\}$  and  $\hat{u}(s, Q, D) = \int_{\bar{D}(s, Q)}^{\min\{\kappa, \bar{D}(p_a, Q)\}} u_Q(Q, D) dF(D)$ .

Recall the buyer's FOC *without an outside option*, from the previous section:

$$\int_{\bar{D}(s, Q)}^{\beta} [u_Q(Q, D) - s] dF(D) = p + s(1 - F(\bar{D}(s, Q))), \quad (3.23)$$

This implies

$$Q^* = [h(p + s(1 - F(\bar{D}(s, Q))))]_+, \quad (3.24)$$

where  $h(x) = \{Q \in \mathbb{R}_+ : \bar{u}(Q; s, D) = x\}$  and  $\bar{u}(Q; s, D) = \int_{\bar{D}(s, Q)}^{\kappa} u_Q(Q, D) dF(D)$ .

The buyer's FOC with an outside supply option contains the additional term  $[p_a(1-F(\bar{D}(p_a, Q)))]$ , which tends to zero for large enough  $p_a$  or large enough  $Q(p, s)$ . Lemma 3.7 below establishes the piecewise nature of the demand curve, which, as we will see, is the source of nonmonotonocities in the optimal price function.

**Lemma 3.7** *The buyer's demand curve in the presence of an outside option is piecewise and given by:*

$$Q^*(p, s) = \left\{ \begin{array}{ll} h(p + s(1 - F(\bar{D}(s, Q))))_+, & \text{if } \bar{D}(p_a, Q(p, s)) \geq \kappa \\ \hat{h}(p + s(1 - F(\bar{D}(s, Q)))) - p_a(1 - F(\bar{D}(p_a, Q)))_+, & \text{otherwise} \end{array} \right\}.$$

where  $h(x) = \{Q \in \mathbb{R}_+ : \bar{u}(Q; s, D) = x\}$  and  $\bar{u}(s, Q, D) = \int_{\bar{D}(s, Q)}^{\kappa} u_Q(Q, D)dF(D)$ , and  $\hat{h}(x) = \{Q \in \mathbb{R}_+ : \hat{u}(Q; s, D) = x\}$  and  $\hat{u}(s, Q, D) = \int_{\bar{D}(s, Q)}^{\min\{\kappa, \bar{D}(p_a, Q)\}} u_Q(Q, D)dF(D)$ .

The number of options contracted for when the buyer has an outside option is greater than the quantity held in the absence of an outside option.

Lemma 3.8 establishes the fact that the ratio governing the buyer's tradeoff between  $p$  and  $s$  is unchanged from the case with an outside option. The fact that this ratio is unchanged has an important implication for the seller's problem, namely that the seller implements the same optimal pricing rules as before *subject to the buyer's IR constraint*.

Lemma 3.8 follow directly from the first-order conditions.

**Lemma 3.8** *The ratio of  $Q_p$  to  $Q_s$  with an outside option is  $\frac{Q_p}{Q_s} = \frac{1}{1-F(\bar{D}(s, Q))} > 1$ .*

The structure of the seller's maximization problem has not changed with the introduction of an outside option for the buyer. We have

$$\max_{p, s} \{pQ(p, s) + (s - b)y + E[(\bar{m} - b)(K - y)]\}.$$

where again  $y(Q, s)$  is the expected number of options that the buyer will exercise, given that he purchases  $Q$ . If  $s < p_a$ , then the seller will exercise all of his options before he considers purchasing additional supply at  $p_a$  (hence  $y(Q, s)$  is given as before). (If the seller sets  $s > p_a$ , then the buyer will neither exercise nor purchase any options!) Furthermore, from Lemma 3.8, the ratio of  $Q_p$  to  $Q_s$  is unchanged.

We obtain the same (implicit) definitions of  $s^*$  and  $p^*$  as in the case with no outside supply option *when the IR constraint is non-binding*, where again  $p^*$  can be written in terms of the option price elasticity of demand and  $s^*$  is always greater than  $\bar{m}$  and is decreasing in the expected number of options to be exercised  $\bar{g}$ .

Letting  $U_{p_a}$  denote the buyer's utility level from relying on his outside option, the buyer's reserve value is

$$U_{p_a}(p_a) = \int_0^{D(p_a)} u(0, D) dF(D) - \int_{D(p_a)}^{\infty} [u(g(p_a, D), D) - p_a g(p_a, D)] dF(D). \quad (3.25)$$

The buyer's IR constraint is given below

$$\phi = U(s, Q) - pQ - U_{p_a}(p_a) \geq 0. \quad (3.26)$$

Theorem 3.6 below characterizes the optimum for the case with a fixed-price outside option for the buyer.

**Theorem 3.6** (*Optimal strike price and option price with an outside option*) Under assumption A1 and A2, the seller sets the optimal strike and option prices in accordance with the following rules

$$s^* = \bar{m} - \frac{\bar{g} - \lambda(\phi_s - \phi_p(1 - F))}{\int_{\bar{D}} g_s f(D) dD}$$

$$p^* = \frac{-(s - \bar{m})(1 - F) + \lambda \frac{\phi_p}{Q_p}}{1 - 1/\varepsilon_p}$$

where  $\lambda$  is the Lagrange multiplier on the buyer's IR constraint,  $\bar{g} = \int_{\bar{D}(s, Q)} g(s, Q) dF(D)$ , and we have written  $(1 - F)$  in place of  $(1 - F(\bar{D}(s, Q)))$ .

We observe that when the buyer's IR constraint is binding, the option price is *always lower* than when the constraint is non-binding. (This from the fact that the denominator and the expression  $-(s - \bar{m})(1 - F)$  are negative – and since the buyer's expected utility and  $Q$  are both decreasing in  $p$ , we have  $\frac{\phi_p}{Q_p} > 0$ .) The strike price is generally lower: the price of the outside option places a restriction on  $s^*$ , namely  $s^* \leq p_a$ , with a strict inequality holding when  $p^* > 0$ . The tighter IR constraint results in a more favorable outcome for the buyer. In the example we develop below, we see, for instance, that the seller sets a lower  $p^*$  (compensating to the degree possible with a higher  $s^*$ , where, again,  $s^* \leq p_a$ ) and the buyer holds a higher number of options  $Q$ . Lemma 10 states this intuitively appealing result.

**Lemma 3.9** *The optimal  $Q^*(p, s)$  is (weakly) decreasing in the price of the outside option.*

### Example: Contract With an Outside Option at a Fixed Price

When the buyer has an outside supply option available at a fixed price  $p_a$ , the seller is clearly constrained to charge a strike price below  $p_a$ . As the seller can get at least  $\bar{m}$  per

unit of supply in expectation by selling on the spot market, there will be no contracting when  $p_a < \bar{m}$ .

The buyer's piecewise demand function is given as:<sup>11</sup>

$$Q(s, p; p_a) = \begin{cases} Q_{gt} = [-\frac{1}{2}s + 1 - \sqrt{p}]_+, & \text{if } 2\sqrt{p} - s < p_a \\ Q_{lt} = [-\frac{1}{4}(4s - s^2 + p_a^2 - 4p_a + 4p)/(s - p_a)]_+, & \text{if } 2\sqrt{p} - s \geq p_a \end{cases}$$

There is a kink in the demand function at the point where  $\min\{\bar{D}(p_a, Q_{gt}), \kappa\} = \kappa = 1$ :

$$\bar{D}(p_a, Q) = Q + \frac{p_a}{2} = \frac{1}{2}s + 1 - \sqrt{p} + \frac{p_a}{2} > 1 \tag{3.27}$$

implies  $2\sqrt{p} - s < p_a$ . For prices satisfying (3.27), the buyer *does not expect to ever rely on his outside option*. We note the dependence of (3.27) on  $p_a$  - it is clear that this is not a restrictive condition for large  $p_a$ .<sup>12</sup>

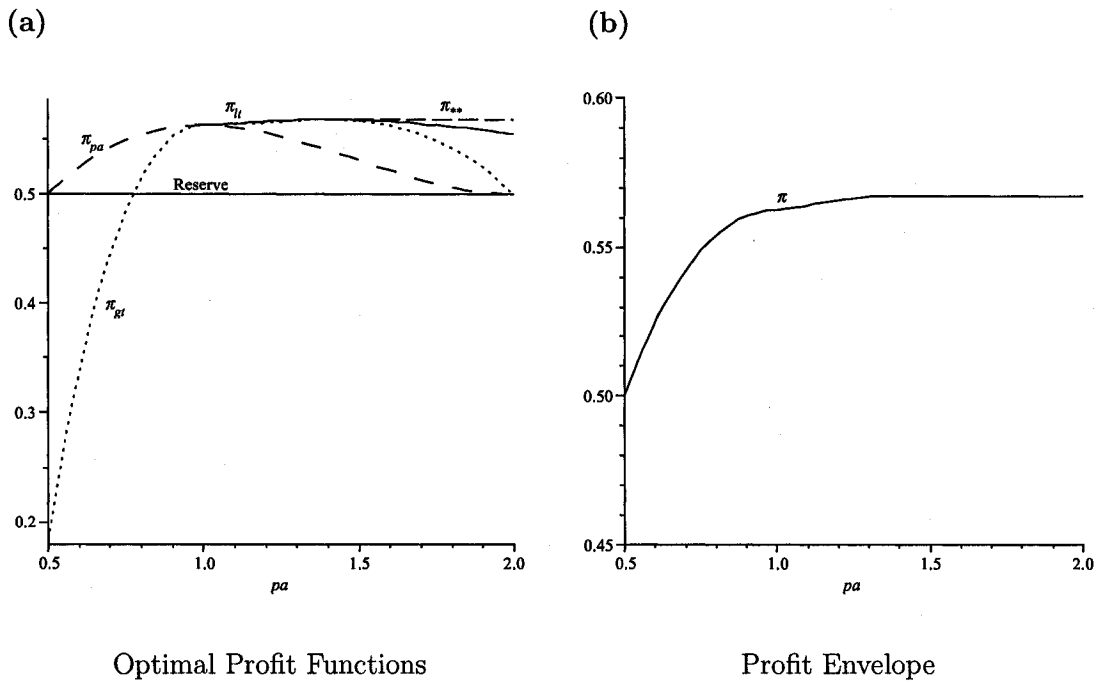


Figure 3.7: Profit Analysis

<sup>11</sup>We note that  $Q_{lt}$  is undefined where  $p_a = s$ . If  $s = p_a$  then it must be  $p = 0$  in order for contracting to take place, in which case the demand function is given by  $Q_{gt}$ .

<sup>12</sup>An explicit solution for the  $p_a$  for which the outside option becomes worthless can be found by setting  $2m - pa/2 = Q_{gt}(s^{**}, p^{**})$ , where  $p^{**}$  and  $s^{**}$  denote the optimal unconstrained prices, i.e., the prices charged without an outside option.

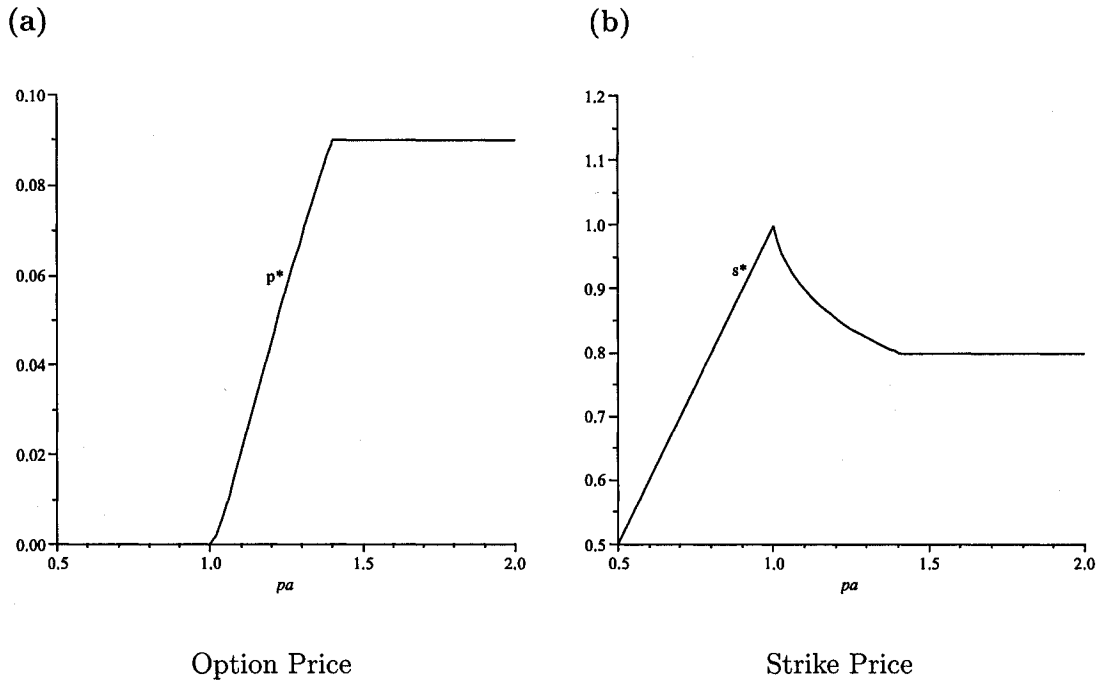


Figure 3.8: Sensitivity of Option Price to the Price of the Outside Supply Option

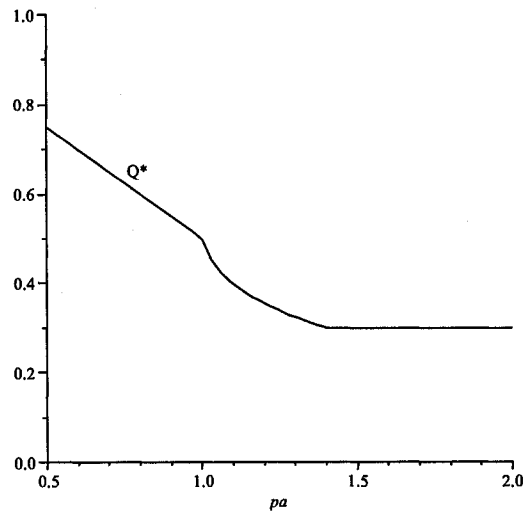


Figure 3.9: Optimal Contract Level

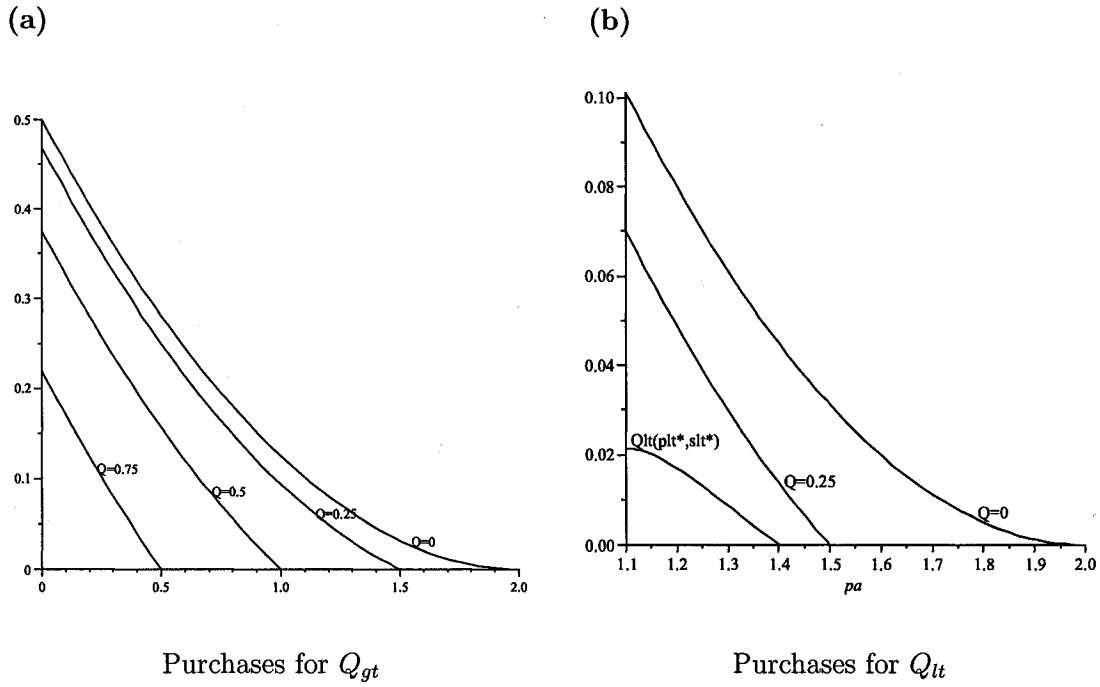


Figure 3.10: Expected Purchases from the Outside Option

We can explicitly solve for the  $p_a$  at which the seller optimally reverts to prices set without an outside option:

$$\bar{D} = Q(p^{**}, s^{**}) + \frac{p_a}{2} = \kappa = 1$$

This implies

$$p_a > \frac{4}{5}\bar{m} + \frac{2}{5} + 2\sqrt{\frac{4}{25} - \frac{4}{25}\bar{m} + \frac{1}{25}\bar{m}^2}$$

For  $\bar{m} = 0.5$ , the inequality is satisfied for  $p_a > 1.4$ . We see in Figure 3.8 that at this point the optimal prices converge to the solution *without an outside option*.

We further observe that for  $p_a < 2$ , the outside option is potentially valuable to the buyer. That is, there are realizations of demand for which the buyer would purchase supply at  $p_a$ . In these cases, the buyer's IR constraint depends on  $p_a$  (as well as  $p$  and  $s$ ):

$$U_{p_a} \leq U(s, Q) - pQ, \tag{3.28}$$

where  $U_{p_a}$  denotes the utility to the buyer of relying solely on the outside supply option at a fixed price  $p_a$ .

Figure 3.7 shows the profit envelopes for prices  $(p_{gt}, s_{gt})$ ,  $(p_{lt}, s_{lt})$ ,  $(0, p_a)$ , and  $(p^*, s^*)$ . The seller optimally charges  $(0, p_a)$  for  $p_a < 1$ . For  $1 \leq p_a < 1.4$ , the optimal prices are  $(p_{lt}, s_{lt})$ . For  $p_a > 1.4$ , optimal prices converge to those charged without an outside option.

Figure 3.8 shows the optimal option price  $p^*$  always below that in the case with no outside option. Clearly the optimal strike price  $s^*$  must remain below the price of the outside supply option in order to satisfy the buyer's IR constraint. We see that the optimal strike price increases in  $p_a$ , reaching an optimal level above that for the case with no outside option. The optimal option price  $p^*$  is at zero for low values of  $p_a$  and gradually increases to the optimal level for the case where there is no outside supply option.

Figure 3.9 shows the optimal quantity of options held  $Q^*$  is decreasing in  $p_a$ , where an increase in  $p_a$  translates to a weakening of the buyer's IR constraint and hence more seller power (and higher prices, where  $Q$  strictly decreasing in  $p$  and  $s$ ). The seller's profit is increasing in  $p_a$ .

### 3.7 Contracting with Joint Spot Market Access

We have assumed up until now that the seller has access to a spot market on which he can offload unclaimed supply and that the buyer has no access to the market. Our motivation for this assumption is existing contracting in nascent water markets, in which there is no spot market as such for the good under contract (water). However, there is seller access to a spot market for a good produced with water as an input (usually a crop). Hence, ignoring the complications of a production function, we arrived at the models in Sections 3.5 and 3.6. The restriction on buyer spot market access also captures the possibility of market-exclusion as a result of infrastructure competition. For instance, in the electricity market the ability to effectively purchase on the spot market is coupled with transmission line access. A buyer without infrastructure access contracts may be essentially excluded from the market.

We now relax the assumption of no buyer access and examine the consequences of joint access to a spot market for the good under contract. Such scenarios will be directly relevant to the water market if the volume of transactions increases to a point where the formation of a spot market is a desirable investment for the participants.

#### 3.7.1 Certain Demand

Wu et al. (2002) treat the case of contracting between a seller and buyer where both have access to the spot market and the buyer's downstream demand  $D$  is known. In their formulation, the seller has imperfect access to the market and can only offload some percentage  $r$  of

his leftover supply onto the spot market in the second period.<sup>13</sup> As discussed in Section 3.4, the seller *only* wants to contract if  $r < 1$ . (For  $r = 1$ , the seller prefers to wait until the spot market price is realized in the second period.) In Wu et al.'s model,  $r$  is described as a “risk factor” that can alternately represent the percentage of the residual output that the seller can sell on the spot market or the probability of finding a willing buyer on the day the output arrives at the market. We assume that the seller has unmitigated access to the spot market:  $r = 1$ .<sup>14</sup>

In order to represent the possible transaction costs associated with purchases through a spot market, i.e., the transportation cost of water, we introduce a coefficient  $\tau$ , which we refer to as the “market friction” coefficient. Our assumption of a constant market friction coefficient  $\tau$  could meet with the same criticism as the use of a constant risk factor in Wu et al.: it assumes a per-unit transaction cost that is independent of the the transaction volume. Current pricing of infrastructure in the California water market is on a constant per-unit basis. However, as the market develops it is plausible that large transactors would negotiate discounted rates.

Under certain demand, the market friction coefficient  $\tau$  in our model naturally plays the same role as the “risk factor”  $r$  – namely, it provides the incentive for contracting to take place. For  $\tau = 0$  in our model, i.e., a frictionless market, there is no contracting under certain demand.

With access to the spot market, the buyer will elect not to exercise any of his options (and instead to buy directly from the spot market) whenever  $m + \tau < s$ , where  $m$  is the realization of the spot market price  $\tilde{m}$  and  $\tau$  is the spot market transaction cost faced by the buyer. For  $m + \tau \geq s$ , the buyer will exercise all  $Q$  options before making *any* purchases from the spot market. If he wants supply in addition to  $Q$ , i.e., if demand realizes such that  $D > \bar{D}(m, Q)$ , then the buyer will procure additional supply from the market at a price  $m$ .

---

<sup>13</sup>In Wu et al. (2002) the percentage of supply that is offloadable, or “risk factor” is denoted as  $m$ . To avoid notational confusion (where in our model  $m$  refers to the realized spot market price), we rename the factor  $r$ .

<sup>14</sup>While the explanations for why a seller might not be able to offload his entire leftover supply on the spot market – a “thin” or disorganized market that makes finding a buyer on the given day difficult or transportation capacity constraints such as may exist in the electricity market (Wu et al., 2002) – are entirely plausible, the representation of these risks as independent of the volume of supply being offloaded is an unsatisfactory assumption. These risks presumably depend on the (expected) volume being offloaded by the seller, which in turn depends on the seller's contracts. If the seller only wants to offload a small quantity on the spot market, we might expect him to face a smaller risk than if he were offloading a large quantity.



**Theorem 3.7** (From Wu et al) *The optimal option and strike prices with joint spot market access and certain downstream demand are given as follows:*

$$p^* = \frac{rE(m-b)_+}{1-1/\varepsilon_p}$$

$$s^* = b$$

*Proof.* See Wu et al. ■

The strike price is simply set to recoup the seller's delivery cost  $b$  and the margin is extracted through the upfront fee. The option price  $p^*$  above takes the form of the familiar inverse elasticity rule for monopoly pricing, discussed earlier:  $p = \frac{MC(D(p))}{1-1/\varepsilon_p}$ , where  $rE(m-b)_+$  is the seller's marginal delivery *plus* opportunity cost. In contrast to the case with no buyer market access, where the seller prices in the elastic part of the demand curve ( $\varepsilon_p < 1$ ), once the buyer gains access to the market the seller prices more aggressively, pricing in the inelastic part of the demand curve. From above,  $\varepsilon_p > 1$ . This intuition behind more aggressive pricing on the part of the seller is as follows: with joint spot market access, the buyer only exercises the options he holds when the spot market price realizes high, i.e., realizes above  $s^*$  – in which case, the seller would have preferred to face the market price. The buyer's unfavorable (to the seller) execution policy leads the seller to charge a higher upfront fee.

### 3.7.2 Uncertain Demand

Now we turn to the case of uncertain demand. When the buyer has access to the spot market and demand is known, he will make a “bang-bang” decision in the second period with regard to his options: he will either exercise all of them (if  $m + \tau \geq s$ ) or none (if  $m + \tau < s$ ). When demand is uncertain, however, the seller expects the buyer to exercise a *percentage* of the options, based on realized demand  $D$ , whenever  $m + \tau \geq s$ . The key changes in the contracting problem for the case with uncertain demand are as follows: (1) the buyer's second-period problem now depends on  $z$  and  $\hat{z}$ , where  $\hat{z} = [g(m + \tau, D) - Q]_+$  is the supply that the buyer purchases from the market, and (2) the number of options that the seller expects the buyer to exercise  $y$  is now proportional to  $P(m \geq s - \tau)$ .

We have

$$z(Q, s, D) = \begin{cases} g(s, D)_+, & \text{if } D < \bar{D} \text{ and } s \leq m + \tau \\ Q, & \text{if } D \geq \bar{D} \text{ and } s \leq m + \tau, \\ 0, & \text{otherwise,} \end{cases}$$

and

$$\hat{z}(Q, m + \tau, D) = \begin{cases} [g(m + \tau, D) - Q]_+, & \text{if } s \leq m + \tau \\ g(m + \tau, D)_+, & \text{otherwise.} \end{cases}$$

where  $\bar{D} = \{D : g(s, D) = Q\}$ .

In the second period, the buyer observes the realized demand  $D$ . Given the exercise price  $s$ , the market fee  $m + \tau$ , and the available quantity of options  $Q$  the buyer solves the following problem in the second period:

$$\max_{z, \hat{z}} \{u(z + \hat{z}, D) - sz - (m + \tau)\hat{z}\}. \quad (3.29)$$

where  $u(z + \hat{z}; D)$  is the buyer's gross payoff when delivering the amount  $z + \hat{z}$ , given demand realization  $D$ .

The buyer's expected utility of contracting for  $Q$  additional units of supply in the first period, obtained by substituting  $z(Q, s, D)$  and  $\hat{z}(Q, m + \tau, D)$  into the objective function in (3.29) and taking the expectation with respect to  $D$ , is therefore

$$U(Q, s, \tau) = E[u(z(Q, s, D) + \hat{z}(Q, \tilde{m} + \tau, D), D) - sz(Q, s, D) - (\tilde{m} + \tau)\hat{z}(Q, \tilde{m} + \tau, D)].$$

As before,  $U(Q, s, \tau)$  is generally increasing, so that at  $p = 0$  the buyer would hold the maximum quantity of options he could ever expect to exercise  $g(s, \kappa)$ .

At a price pair  $(p, s)$  the buyer's first-period problem is to maximize his expected payoff

$$\max_{Q \geq 0} \{\hat{U}(Q, s, \tau) - pQ\}.$$

The buyer's objective function is given as

$$\begin{aligned} \max_Q \{U(Q, s, \tau) - pQ\} &= \int_0^{s-\tau} \int_0^{D(\tilde{m}+\tau)} u(0, D) dF(D) dF(m) \\ &+ \int_0^{s-\tau} \int_{D(m+\tau)}^\kappa [u(g(m+\tau, D), D) - (m+\tau)g(m+\tau, D)] dF(D) dF(m) \\ &+ P(m \geq s - \tau) \left[ \int_0^{D(s)} u(0, D) dF(D) + \int_{D(s)}^{\bar{D}(s, Q)} [V(g(s, D), D) - sg(s, D)] dF(D) \right] \\ &+ \int_{s-\tau}^\beta \int_{\bar{D}(s, Q)}^{\bar{D}(m+\tau, Q)} [u(Q, D) - sQ] dF(D) dF(m) \\ &+ \int_{s-\tau}^\beta \int_{\bar{D}(m+\tau, Q)}^\kappa [u(g(m+\tau, D), D) - sQ - (m+\tau)(g(m+\tau, D) - Q)] dF(D) dF(m) - pQ \end{aligned}$$

**Proposition 3.3** *The buyer's optimal contracting decision is given as*

$$Q^*(p, s) = \left[ h \left( p + s \left[ \frac{1 - C}{1 - \hat{F}_m(s - \tau)} \right] - (m + \tau) \left[ \frac{1 - B}{1 - \hat{F}_m(s - \tau)} \right] \right) \right]_+$$

where  $h(x) = \{Q \in \mathbb{R}_+ : \bar{u}(Q; s, D) = x\}$  and  $\bar{u}(Q; s, D) = \int_{s-\tau}^{\beta} \int_{\bar{D}(s, Q)}^{\bar{D}(m+\tau, Q)} [u_Q(Q, D)] dF(D) dF(m)$  and  $C = \hat{F}(\gamma, \bar{D}(s, Q)) - \hat{F}(s - \tau, \bar{D}(s, Q))$ ,  $B = \hat{F}(\gamma, \bar{D}(m + \tau, Q)) - \hat{F}(s - \tau, \bar{D}(m + \tau, Q))$ , and  $\hat{F}$  is the joint cdf of  $\tilde{m}$  and  $\tilde{D}$  and  $\hat{F}_m$  is the cdf of  $\tilde{m}$  and  $\gamma$  is the upper bound of support of the market price  $\tilde{m}$ .

The seller's problem has also changed. The expectation of the spot market price and the number of options exercised are no longer independent – if the spot market price realizes high ( $m > s - \tau$ ) then no options are exercised. Otherwise, the seller expects  $y$  options to be exercised. The seller's problem is given as

$$\max_{p, s} \{pQ + P(\tilde{m} \geq s - \tau)(s - b)y + E[(\tilde{m} - b)(K - y) | \tilde{m} \geq s - \tau]P(\tilde{m} \geq s - \tau) + E[(\tilde{m} - b)K | \tilde{m} < s - \tau]P(\tilde{m} < s - \tau)\}$$

which simplifies to

$$\max_{p, s} \{pQ + P(\tilde{m} \geq s - \tau)(s - E[\tilde{m} | \tilde{m} \geq s - \tau])y + (\tilde{m} - b)K\}, \quad (3.30)$$

where, as before

$$y_p = Q_p(1 - F(\bar{D}))$$

$$y_s = Q_s(1 - F(\bar{D})) + \int_{\bar{D}(s)}^{\bar{D}(s, Q)} g_s(s, D) f(D) dD.$$

although our expressions for  $Q_p$  and  $Q_s$  have clearly changed.

The seller's FOCs characterize the optimal  $p$  and  $s$  when the seller's payoff function is jointly concave. This characterization is given in Theorem 3.8 below, the proof of which is given in the Appendix.

**Theorem 3.8** *(Optimal option price and strike price with joint spot market access) Under assumptions A1 and A2, the seller optimally sets prices according to the following rules:*

$$s^* = \frac{\hat{m}P(m \geq s - \tau)\hat{g}_s - y\Gamma + Q\Upsilon}{(P_s + f(s - \tau))y + P(m \geq s - \tau)\hat{g}_s}$$

$$p^* = \frac{-(s - \hat{m})(1 - F)P(m \geq s - \tau)}{1 - 1/\varepsilon_p}$$

where  $\hat{m}$  is the conditional mean of the spot market price, conditional on the fact that  $\tilde{m}$  realizes above  $s - \tau$ ,  $\hat{g}_s = \int_{\bar{D}(s)}^{\bar{D}(s, Q)} g_s dF(D)$ ,  $\Gamma = P(m \geq s - \tau) - \tau f(s - \tau)$ , and  $\Upsilon = \frac{Q_s}{Q_p}$ .

We observe that the aggressiveness of the seller's pricing of the option, as indicated by whether he prices in the elastic or inelastic part of the demand curve, depends critically on the strike price that he is able to charge. This strike price in turn depends on both the probability that the buyer will call all  $Q$  options ( $1 - F(\bar{D}(s, Q))$ ) and the probability that the market price realizes high  $P(m \geq s - \tau)$  and is *increasing* in the former and *decreasing* in the latter.

### 3.7.3 Example: Contracting with Joint Spot Market Access

As in the case with an outside option at a fixed price, whether or not the buyer ever expects to rely on his outside option (the spot market) to meet supply depends on the number of options he holds  $Q(p, s)$  and the exercise price  $s$ . The buyer's demand curve has two kinks: (1) if prices are set such that  $s^* \leq \tau$ , this implies  $Q + \frac{s-\tau}{2} < 2\bar{m}$ , since  $Q < 2\bar{m}$  necessarily), then the buyer will *always* exercise all options before going to the market, irrespective of the realization of  $\tilde{m}$ , (2) if prices are set such that  $s^* > \tau$  and  $Q(p^*, s^*) + \frac{s-\tau}{2} > 2\bar{m}$ , then the buyer will use the market to fill demand whenever  $\tilde{m}$  realizes below  $s - \tau$  however, he will never fill excess demand from the spot market *when* the spot market price  $\tilde{m}$  realizes above  $s - \tau$ , and, finally, (3) if prices are set such that  $s^* > \tau$  and  $Q(p^*, s^*) + \frac{s-\tau}{2} < 2\bar{m}$ , then as in (2) the seller will choose the market over the option contract whenever  $\tilde{m}$  realizes below  $s - \tau$  and he may also fill excess demand from the spot market when the spot market price  $\tilde{m}$  realizes above  $s - \tau$  and after he has exhausted all of his options. It is clear that prices set as described in (3) leave the buyer holding less options  $Q$  and hence present the possibility of reliance on the spot market in the case of high demand realization, i.e., cases where the buyer needs more than  $Q$  options.

The buyer's piecewise demand curve is given as follows:

$$Q(s, p; p_a) = \begin{cases} Q_1 & \text{if } s < 4\bar{m} - 2Q(s, p) + \tau < \tau \\ Q_2, & \text{if } s > 4\bar{m} - 2Q(s, p) + \tau > \tau, \\ Q_3, & \text{if } \tau < s < 4\bar{m} - 2Q(s, p) + \tau, \end{cases}$$

where we determine  $Q_1$ ,  $Q_2$ , and  $Q_3$  numerically for our example.

In Figures 3.11 and 3.12 we observe the following behavior: the seller optimally charges  $s^* \geq \tau$ , which means that there is always the possibility that the buyer will meet some excess demand on the spot market when  $\tilde{m}$  realizes above  $s - \tau$ . (And clearly when  $\tilde{m}$  realizes below  $s - \tau$ , the buyer fills all demand on the spot market.) The number of options demanded is lower than when the buyer had no market access, and the seller charges a higher upfront

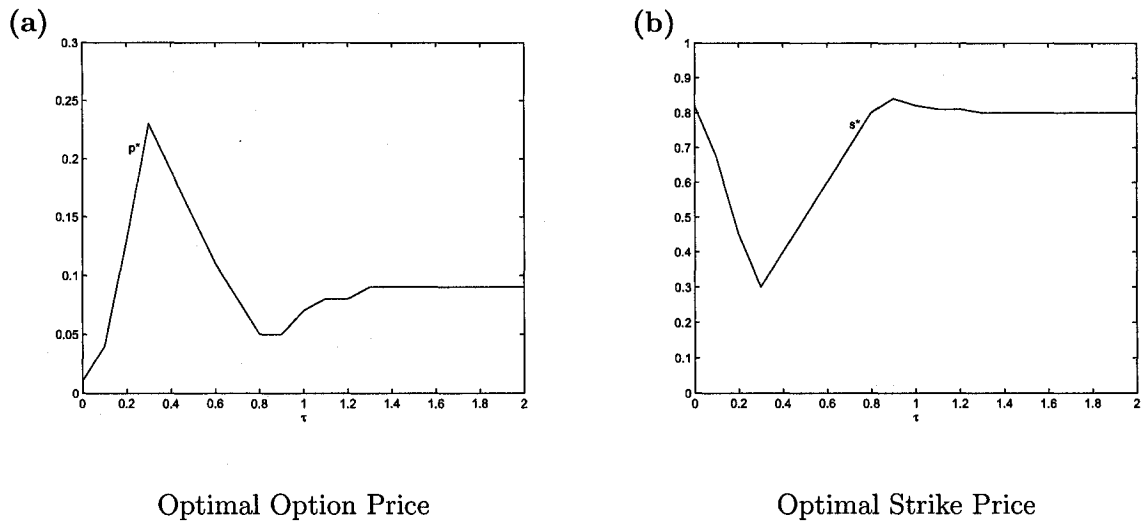
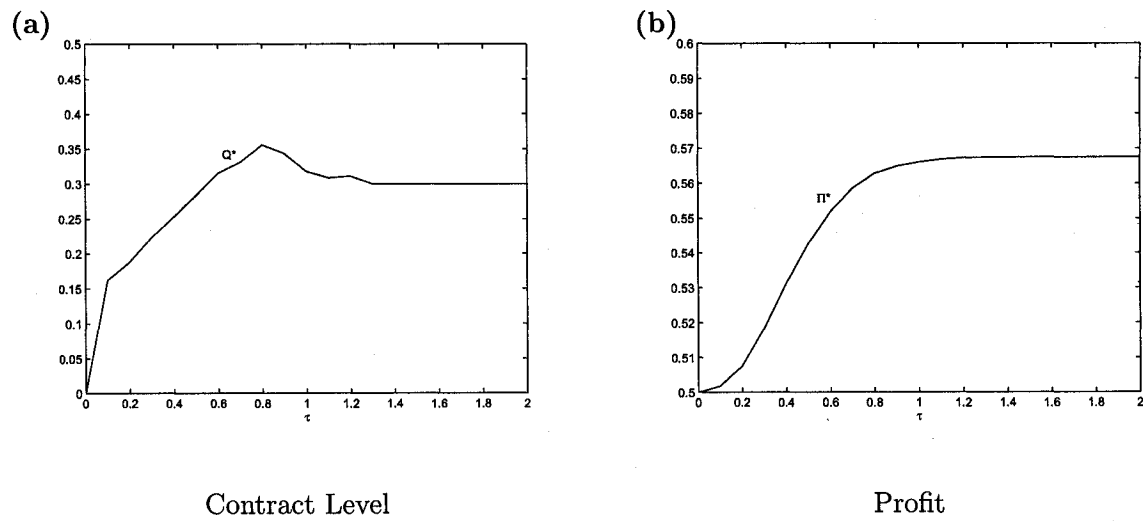


Figure 3.11: Optimal Prices with Joint Spot Market Access

Figure 3.12: Contract Level and Profit Sensitivity to the Market Friction Coefficient  $\tau$

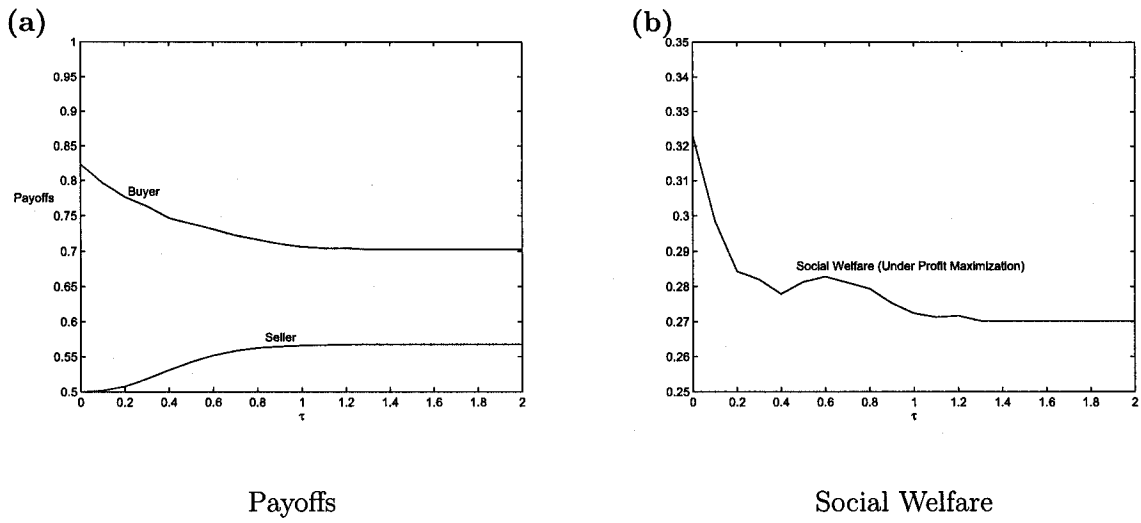


Figure 3.13: Payoff and Social Welfare Sensitivity to the Market Friction Coefficient  $\tau$  (Under Profit Maximization)

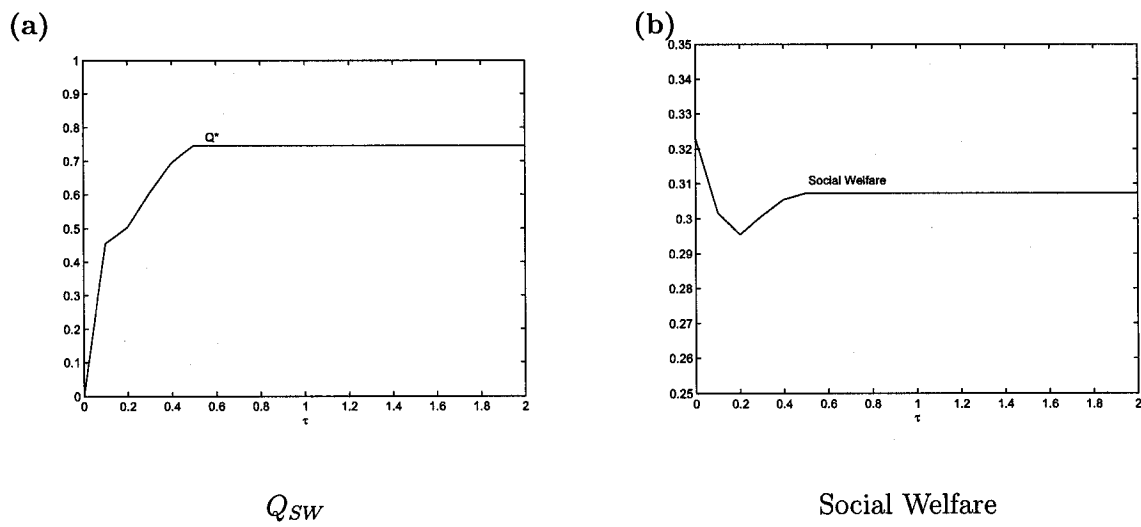
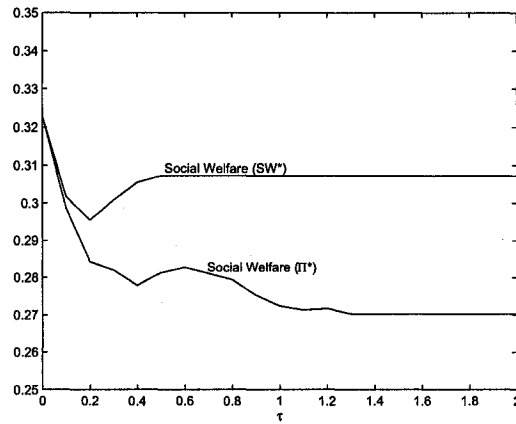
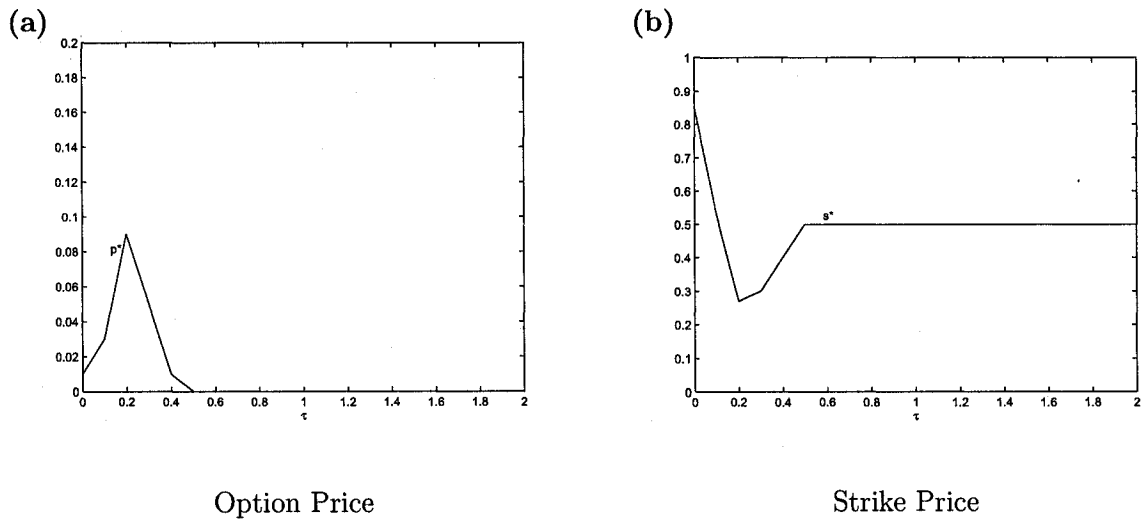


Figure 3.14: Order Level and Social Welfare Sensitivity to the Market Friction Coefficient  $\tau$  (Under Social Welfare Maximization)



Social Welfare

Figure 3.15: Social Welfare Comparison Under Profit Maximization vs. Social Welfare Maximization



Option Price

Strike Price

Figure 3.16: Social Welfare Optimal Prices

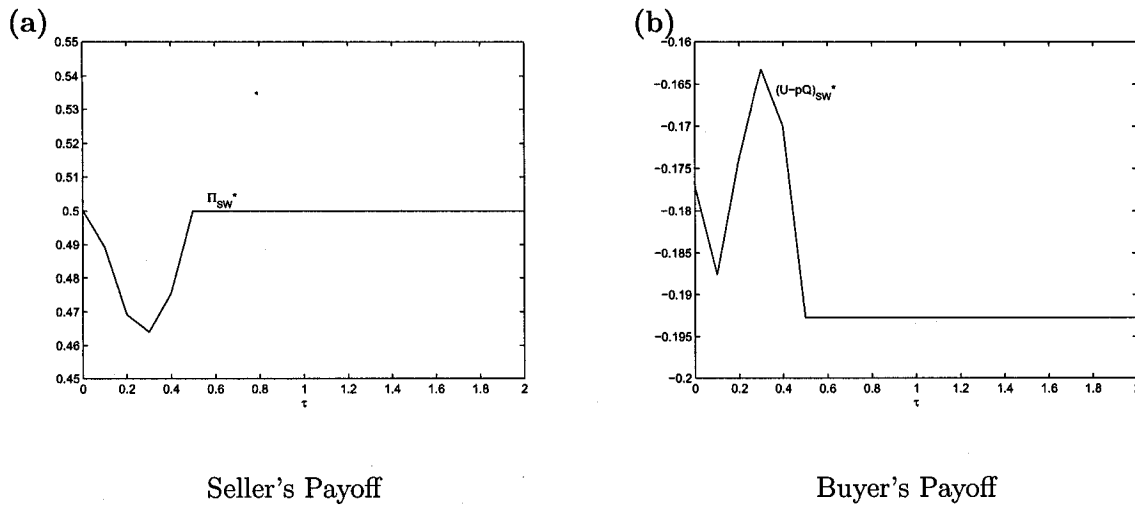


Figure 3.17: Payoffs Under Social Welfare Maximization

cost. The optimal option price  $p^*$  is increasing in  $\tau$ , while the optimal strike price  $s^*$  is below the optimal no-option-case value, for low values of  $\tau$ . Specifically, we observe that  $s^* < m$  in these cases. The optimal number of options purchased  $Q^*$  is increasing in  $\tau$ . The seller's profit is also increasing in  $\tau$ , since an increase in  $\tau$  corresponds to a slackening of the buyer's individual rationality constraint, thereby allowing the seller to increase prices. Here  $m = 0.5$ .

There is a market transaction cost  $\tau$  for which access to the market by the buyer is effectively meaningless, i.e., a  $\tau$  that makes transacting on the market too costly for any realization of demand. In which case, the seller will price as he did in the case with no access. In Figure 3.11 we see that this convergence point occurs at  $\tau = 1$  (for  $\bar{m} = 0.5$ ). The value of  $\tau$  at which the outside option becomes worthless is lower than that of  $p_a$ , since, in expectation, the buyer expects to pay more than  $\tau$  per unit on the spot market. (Recall that the convergence point above was at  $p_a = 1.4$ .)

In Figure 3.13, we see that the seller's profit is increasing in  $\tau$  and the buyer's payoff is decreasing. Social welfare is decreasing in  $\tau$  under the seller-set profit maximizing contract and exhibits a "dip" (and hence is increasing in  $\tau$  for small values). In Figure 3.14, we see that the socially optimal  $Q$  is always higher than that in the profit-maximizing case. The social welfare again dips for low values of  $\tau$ , when charging a high  $s$  will result in too few options being purchased, from the social planner's perspective, but for  $s < \tau$ , the probability that the buyer will exercise the options when the demand realizes high is non-zero and hence cuts into the seller's expected value. Figure 3.15 compares the "dips" in social welfare associated



with low values of the market friction coefficient,  $\tau$  for cases of profit-maximization and social-welfare maximization. This “dip” can be considered to some degree an artifact of the option structure and, in particular, the fixed strike price. If the social planner indexed the option to the market price, then social welfare would be a smoothly decreasing function of  $\tau$ . In the case of a fixed strike price, however, we have the counter-intuitive result that social welfare is increasing in the market friction coefficient, for low values of  $\tau$ . Figure 3.16 shows that, in an attempt to offset this affect, the social planner actually charges an upfront fee for low values of  $\tau$  in the standard (unindexed) contract. Figure 3.17 shows the seller’s and buyer’s payoffs, respectively, under social welfare maximization.

## 3.8 Application to the California Water Market

### 3.8.1 Background

In 2003, the Metropolitan Water District of Southern California (MWD) initiated option contracting in the California water market. MWD has entered into option contracts with a number of Sacramento Valley farmers over the last several years. As discussed in Chapter 1, these contracts are a new development in the state’s nascent water market, and they represent a shift from an earlier focus on permanent water rights exchanges to one on temporary transfers. If successful, routine option contracting between large urban water districts and agricultural entities could promote a more active water market. Already, MWD is no longer the only buyer. The San Diego County Water Authority (SDCWA) entered the contract market for water options with purchases from two northern California irrigation districts in 2008.<sup>15</sup>

Temporary transfers as achieved under option contracts can significantly reduce market friction, lowering transaction cost and institutional resistance to water transfers. Permanent water transfers in the California water market have associated with them both high transaction cost and considerable institutional resistance (Hanak 2003; Hanak 2005; Hanak and Howitt 2005). Members of farming communities whose livelihoods depend on crop cultivation (e.g., equipment sales- and repairmen, marketers, and millers), and those who do not stand to gain financially from the sale of water, may seek political means to block water transfers (Thompson 1993). In agricultural water districts where all members have voting rights, as opposed to just landowners, there may be sufficient political power to block trans-

---

<sup>15</sup>The 2005 option contracts signed by MWD included several other State Water Project contractors, although MWD took the bulk of the water.

fers. Even in districts with landowner voting rights, there is an incentive for farmers who plan to purchase extra water to block intradistrict transfers, thereby keeping the interdistrict prices low. The due diligence requirements, which include an assessment of third-party and community impacts and a full environmental review under the California Environmental Quality Act (CEQA), contribute to the high transaction cost.

In lieu of permanent or long-term transfers, temporary water transfers can avert costly supply shortfalls. Recognizing the importance of temporary transfers in managing short-term supply risk, the Department of Water Resources (DWR) instituted the Emergency Drought Water Bank (EDWB) during a critically dry period in the early 1990s. In its first year of operation, 1991, DWR purchased over 400,000 af of water from agricultural districts for \$125/af and then sold the water to urban agencies for \$175/af, the price mark-up intended to cover the transaction cost.<sup>16</sup> The EDWB was operated again in 1992 and 1994, and then more recently in 2001, 2002, and 2003.<sup>17</sup> This centralized operation has the advantage of buyer aggregation, as well as reliable access to infrastructure, of which DWR retains control. In the long-term, however, it is doubtful that EDWB operations can scale to adequately and efficiently meet the needs of the hundreds of individual water districts in California that could reasonably incorporate temporary transfers into their water-management strategies.

Option contracts provide an alternative means of implementing temporary transfers. There have been 15 short-term option contracts signed to date, in 2003, 2005, and again in 2008. In 2003, MWD signed option contracts with 11 separate irrigation districts in the Sacramento Valley for a total volume of 146,230 af.<sup>18</sup> Again in 2005, a dry year, MWD repeated contracting with three of the irrigation districts in Sacramento. In 2008, SDCWA entered the option market for the first time, signing two separate contracts. In contrast to the long-term contract, which took several years to negotiate and millions of dollars in environmental and community mitigation fees, the option contracts were negotiated and executed in the course of several months with limited administrative overhead.

In five years, water option prices have risen from \$10/af base and \$90/af strike (in 2003) to \$50/af base and \$200/af strike (in 2008). One interpretation of the price increase is an

---

<sup>16</sup>Just slightly over half of the water purchased was actually sold that year; some was stored until the following year and sold then.

<sup>17</sup>The EDWB has been rechristened the Dry Year Water Purchase Program.

<sup>18</sup>In 2003, MWD signed contracts with Glenn-Colusa Irrigation District, Western Canal Water District, Richvale Irrigation District, Reclamation District 108, River Garden Farms, Natomas Central Mutual Water Company, Meridian Farms Mutual Water Company, Pelger Mutual Water Company, Pleasant Grove-Verona Mutual Water Company, Sutter Mutual Water Company, and Placer County Water Agency. Only the options from the contract with the Placer County Water Agency were ultimately not called.

increase in seller bargaining power. The option pricing model developed in this chapter assigns the price-setting power to the seller. In reality, the contracts are negotiated between the buyer and the seller. This fact notwithstanding, an examination of seller-optimal pricing provides a benchmark for comparison between current prices and those in a seller-centric market. A future increase in seller power, as foreshadowed by the entrance of additional buyers into the contracting market and both the learning and information dissemination accompanying past sales, may ultimately make this modelling assumption more representative.

### 3.8.2 Contract Parameters

The contracts MWD signed with agricultural water districts specified the fallowing of rice acreage in exchange for a strike payment per acre-foot of conserved water. The conserved water to be made available for transfer to MWD was calculated based on consumptive crop use, defined as the estimated volume of water absorbed by the plant and evaporated from the plant and land surface (referred to technically as evapotranspirative use) and distinct from the total amount of water applied to the crop.<sup>19</sup> The seller's expected value from contracting is a combination of his revenue from the sale of the options and his expected revenue from the exercise of the options and the offload of excess supply onto the commodity market. The seller's reserve value is the expected revenue from applying the water to his rice crop and offloading the entire supply onto the market. Table 1 reports the marginal product of water for rice based on the production cost per acre of rice from recent USDA farm surveys, the yield per acre of cultivated rice, as estimated by the USDA, and the consumptive-use calculation provided by the California Department of Water Resources (DWR) (USDA 2004; USDA 2008; DWR 2008).<sup>20</sup> The marginal product of an acre-foot of water for rice in 2008, for example, with a price of \$13.5/cwt on the market and a \$2.35/cwt subsidy, was \$34. The range of rice prices reported in Table 1 are based on market prices reported by the Farm Service for 2008. Regarding price expectations, farmers have a choice each year to sign a contract with a rice marketer even before planting, which locks in a payment per hundred-weight (cwt) of rice. Alternatively, farmers can join a marketing cooperative and receive monthly installments from coop sales. A third option is to wait until harvest and then sell the rice directly to a marketer at the current price or pay storage fees and offload the rice onto the market at a later date.

The buyer's value under contracting, as developed in the model in Section 2.2, depends on

---

<sup>19</sup>The difference in volume between consumptive use and applied use is the volume of water that returns to the environment in the form of runoff or recharge.

<sup>20</sup>The average cost per acre of rice cultivation for 2003 and 2005 is based on the reported cost in 2000.

Parameter	Unit	Estimate/Range
Price	\$/cwt	13.5 ... 19
Subsidy	\$/cwt	2
Yield	cwt/acre	71.5
Revenue	\$/acre	858 ... 1,753
Average Cost	\$/acre	833
Profit	\$/acre	133 ... 526
Water Use	af/acre	3
MP of Water	\$/af	40 ... 159

Table 3.1: Rice Production Data

four factors: (1) the distribution of excess demand, (2) the shortage cost of water, (3) the conveyance cost and losses, and (4) the probability of non-conveyance. The excess demand distribution used by the contracting agencies is not reported in the contracts signed to date. The assumption in place here is that excess demand is uniformly distributed, and that in the past buyers held enough options to fully meet their excess demand. As discussed in Section 3.8.3 below, past contract prices have been highly favorable to the buyers, a fact that supports the full-coverage assumption.

An appropriate estimate for the shortage cost of water depends on whether the shortage will in fact be covered by the water agency, through a secondary (and presumably more expensive) supply source, or whether it will be passed on directly to consumers. SDCWA has a policy to meet excess retail demand through additional water purchases at a “penalty rate” of \$1,347/af (from MWD). The penalty rate, well above the regular Tier 1 rate that MWD charges its member agencies for contracted supply, is in effect for supply requests above the contracted level.<sup>21</sup> The MWD penalty rate is a proxy for the market rate for an extra acre-foot of supply. Other possible proxies for the shortage cost might include the cost per acre-foot of supply replacement via an alternative technology, such as desalination, currently estimated at \$800-\$900/af. While SDCWA, has plans for construction of a desalination plant, that the facility is not currently online makes it an inappropriate cost estimate for short-term planning purposes. If rather than cover the excess demand, SDCWA planned to pass the shortage on to consumers, the estimate for the shortage cost becomes more complex. The estimate must now take into account the loss in consumer welfare, as well as the cost of a “water outage,” which might include business interruption and landscape

<sup>21</sup>The penalty rate is technically only in effect once the member agency exceeds contracted deliveries by more than 10%.

losses. A comprehensive study in 1993, commissioned by SDCWA and conducted by the economic consulting firm CIC, Inc., reported sectoral shortage costs that, when aggregated, totaled \$5,554/af. There is a need for additional, and updated, investigation in this area. A constant shortage cost implies that the buyer will fully cover demand when he has sufficient options to do so. In other words, the buyer exercises all options up to the realized demand level  $D$ , assuming conveyance is possible.

Year	Access Charge	Stewardship Charge	System Power Charge	Total
2003	\$141	\$23	\$89	\$253
2005	\$152	\$25	\$81	\$258
2008	\$145	\$25	\$110	\$278

Table 3.2: MWD's Unbundled Wheeling Charges (per af).

High energy prices. The marginal cost of water conveyance depends on the price of electricity. It requires 2,900 kWh to transport an acre-foot of water from northern California to Southern California, including pumping over the Tehachapi Mountains.<sup>2</sup> The price of electricity for conveyance, as negotiated under Department of Water Resource contracts, has been stable over the past four years at \$0.05/kWh. A rise in the price of electricity would make water supply in general, and transfers in particular, more expensive. (All water supply relies on energy as an input at some stage, from treatment to conveyance to distribution.) Desalination, which is currently more energy-intensive than water transfers, would become less attractive under a scenario of rising electricity costs. Permanent transfers that require the buyer to pay the conveyance fee may also be less attractive when rising electricity costs are taken into consideration. Temporary transfers mitigate the risk associated with both of these alternatives by providing the flexibility to (1) avoid or postpone desalination investment - or temporarily cease operations if the investment is in place and (2) avoid the upfront investment in securing a permanent transfer that may become uneconomical under rising energy costs.

There is a high probability of non-conveyance. SDCWA reports assessing a 50% probability that the San Francisco Bay Delta pumping plant will be over-capacitated, restricting the movement of water north-south. If conveyance is possible, there is a standard charge per-acre-foot conveyed. The unbundled MWD rate consists of three parts – the first two are fees levied by the State Water Project for use of the infrastructure, which they control, and the third is an environmental surcharge levied by MWD. These three charges are reported in Table 2, for the relevant years. North-south conveyance requires an average of 2,200 kWh of electricity per acre-foot to pump the water south over the Tehachapi Mountains via a series

of lifting stations. There is also typically a system access charge, a rate designed to recoup the infrastructure cost. In 2008, MWD charged a conveyance, or “wheeling,” rate of \$278/af: a \$145/af system access rate, a \$110/af system power rate, and an environmental surcharge of \$25/af. The current system power rate of \$110/af reflects an electricity cost of \$0.05/kWh. The power contracts, negotiated by the Department of Water Resources, are for large volumes of power and have been price-stable in the past. Finally, there is an associated conveyance loss. The 2003, 2005, and 2008 contracts all specified that the conveyance losses associated with transferring water would be borne by the buyer, as would the cost of conveyance. The transfer of water North to South via the San Francisco Bay Delta results in significant conveyance losses due to the need for retention of a percentage of transferred water to meet Delta water quality standards. (This water is often referred to as “carriage water.”) MWD estimates carriage water losses at 20%. The water to be transferred is pumped from the south conveyance point of the Delta into the California State Water Project aqueduct. There is a significant probability that in a given year there will be insufficient capacity available in the conveyance system to pump the water. SDCWA estimated the probability of non-conveyance, or infrastructure failure, as 50% (SDCWA 2008).<sup>22</sup>

### 3.8.3 Welfare Gains Under Option Contracting

A calculation of the magnitude of expected social welfare gains under past contracting requires both knowledge of the contract structure and estimates of the reserve values of the buyer and seller. The seller’s reserve value can be estimated as a function of the then-expected spot market price for the crop that he would otherwise cultivate with his water. As farmers’ expectations for the 2003, 2005, and 2008 commodity prices are unrecorded, the welfare estimates reported in Table 3 are based on the historical prices for rice. The base case corresponds to a calculation of the reserve value using the *actual* commodity price that year, as reported by the Food and Agriculture Organization of the United Nations (FAO). The low-case reflects an underestimate of the actual price by 25% – the high-case an overestimate. As the expected commodity price increases, boosting the reserve values, the seller’s surplus from trade is slowly eroded. The estimated social welfare gains are highest in the

---

<sup>22</sup>The risk of non-conveyance, or non-delivery, was purely an infrastructure risk. The contracts specified that the water to be made available through following arrangements was to be water to which pre-1914 water rights were held. Under the California Water Code, pre-1914 water rights are the most senior water rights. As such, they are considered firm rights: there is always enough water in the system to ensure that priority rights holders receive their allocation, with junior rights holds entitled to the rest of the annual supply, in order of priority.

low-price case. In 2003, the average price of rice for U.S. producers was \$6.25; in 2005, it was \$7.26 (FAO 2008). The 2008 prices are based on the actual market prices. There is also a federal subsidy on rice production of \$2.35/cwt that is included in the calculations.

	2003			2005			2008		
Surplus	Low	Base	High	Low	Base	High	Low	Base	High
Buyer (%)	90.6	92.6	94.6	90	90.3	95.2	66.9	70	73.4
Seller (%)	9.4	7.4	5.7	10	9.7	4.8	33.1	30	26.6
Total (\$)	34.7M	34M	33.2M	29.1M	29M	27.5M	4.81M	4.59M	4.37M

Table 3.3: Surplus Division and Social Welfare Gain with Option Contracts

Three additional assumptions are embedded in the calculations in Table 3. The first is with regard to the excess demand distribution. As discussed above, the assumption used here is that of uniformly distributed excess demand with the mean equal to half the actual contract size.<sup>23</sup> The remaining two assumptions are with regard to shortage cost and conveyance. The shortage cost is assumed to be constant and equal to the penalty rate for additional supply (\$1,347) discussed above. The probability of non-conveyance is assumed to be 50%, based on SDCWA's assessment. Finally, the cost of conveyance is that charged by MWD in each of the years, respectively, as reported in Table 2. Under these assumptions, the social welfare gains in 2003 and 2005 are estimated to be upwards of \$33M and \$27M, respectively. The anticipated gains in 2008 are significantly smaller due to a reduction in the size of the contracts.

### 3.8.4 Price Trends

It appears that past contract prices have been highly favorable to the buyer. There are three likely explanations for this fact. The first explanation is that of buyer bargaining power. In so much as division of surplus is an indication of bargaining power, the calculations in Table 3 suggest that MWD was in the stronger position. The second explanation is one of informational asymmetry. If sellers were unaware of the true value of the options to the buyer, they may have settled for lower prices. Lastly, changes in the modelling assumptions regarding the actual shortage cost of water, the distribution of excess demand faced by the buyer, the seller's expected commodity price, and both the cost and probability of

<sup>23</sup>For a uniform demand distribution, the implication of the assumption is that the buyer has contracted to cover the uppermost demand realization.

conveyance, as well as the one-shot nature of contracting, would alter estimates offered in Table 3.

While the sellers collectively had an estimated expected surplus of between \$1.8M and \$3.3M in 2003, MWD appropriated nearly 90% of the total surplus, or \$31.4M. There are two important factors worth noting with regard to bargaining power. The first is that MWD initiated contracting, contacting farmers regarding the sale of options on their water. As such, MWD was also in the position to make an initial offer. If MWD credibly conveyed the threat to leave the negotiation at any point, the sellers may have faced an “accept/reject” offer with “accept” as the higher-value choice, yielding a payoff above their reserve values.

A second factor acknowledges the political power held by the institutions involved in the transactions. MWD has 26 Southern California member agencies, including the utilities of the state’s two largest cities, Los Angeles and San Diego. The political power in these two constituencies alone is considerable. Wherever there exists an imbalance of political power, there also exists an implicit threat of political recourse. In California, the implicit threat is that of an involuntary reallocation of water. Agricultural water rights could be challenged on efficiency grounds, as they were in IID, or water could be confiscated based on an argument for the primacy of human (urban) use, as has been done in other states (notably New Mexico).<sup>24</sup>

Related to the issue of buyer bargaining power is that of informational asymmetry. While information regarding water management in California is ostensibly public information – urban and agricultural water providers in the state are all public intermediaries – an urban water agency such as MWD has some degree of operational flexibility, making the quantification of both excess demand and the true shortage cost of water difficult for an outside entity. For example, an urban water agency may elect to draw down storage in a given year as opposed to relying on outside water purchases to meet excess demand. There may also be additional transfer options from competing sellers. Finally, there is the possibility of rationing, with the associated implied consumer surplus losses and implementation costs. If MWD successfully convinced sellers that it had alternative low-cost supply options available – or if in reality this was the case – the low contract prices may actually represent a more even division of surplus than reported here.

The trend in contract prices is upwards. The 2003 contract set a \$10/af base and a strike indexed to the year type of \$90/af for normal or wet years and \$115/af for a critically dry

---

<sup>24</sup>Under the California Water Code, retention of a water right requires “beneficial use” of the water. Water rights can therefore be confiscated, or amended, if the water associated with the right is being wastefully applied. This particularly applies to agricultural users, who may be using outmoded irrigation techniques.



Year	Buyer	Contract Type	Volume (af)	Execution (af)
2003	MWD	Options	146,230	126,230
2005	MWD	Options	112,495	0
2005	Other SWPS	Options	14,780	0
2008	MWD	Direct Purchase	28,674	28,674
2008	Other SWPs	Direct Purchase	13,493	13,493
2008	SDCWA	Options	24,038	24,038

Table 3.4: Historical Contract Types and Transfer Volumes.

years.<sup>25</sup> The base price for the 2005 contract was the same; the strike price was higher. The 2005 contract also had a more complex structure, with an extension clause.<sup>26</sup> The indexing feature from the 2003 contract was removed, and instead the strike price was increased. Under the extension clause, the buyer could pay an additional \$10/af to extend the option from April 1 to April 16. The option could be extended a second time from April 16 to May 2, for an additional \$20/af. There was a minimum number of options to be extended, e.g., 40,000 out of the 80,000 held in the GCID. The strike price was set at \$125 minus any option fees already paid, plus an additional land preparation charge of \$10/af between April 16 and May 2 and \$20 after May 2. For example, the strike price between April 16 and May 2 was \$105: the \$125 minus the option payment of \$30 plus the land preparation charge of \$10. The prices in the 2008 SDCWA contracts were again higher than those in the 2003 and 2005 contracts signed by MWD. Part of the explanation for the increase in price owes to an increase in the seller's reserve value. The price of rice on the global exchange has increased five-fold over the last year, as illustrated in Figure 5. At a contract offer of \$10/af upfront and \$90/af upon exercise, a seller would prefer to cultivate rice than to contract his water. Hence, prices necessarily adjusted upward to meet the seller's rationality constraint. Nonetheless, a comparison of the sellers' expected surplus in 2003 and 2005 to that under the 2008 contracts lends credence to the argument that there has also been an increase in seller power. The percent of the expected social surplus

<sup>25</sup>The CA Department of Water Resources uses a five-point scale to designate each water year – critical is the driest designation.

<sup>26</sup>The 2005 MWD contracts contained several changes from the 2003 contracts. First, they pushed the execution date back considerably. The 2003 contract specified an execution date of February 15. A later execution date is favorable to the buyer, as more information regarding storage level and rainfall conditions becomes available later in the season, allowing a more accurate calculation of true excess demand. In California, where the spring rains are highly variable, the buyer is likely to update his forecast significantly between February and May.

received by the sellers increased from less than 10% in 2003 and 2005 to above 27% in 2008. The increase in seller power may be attributed in part to buyer competition. Not only did SDCWA enter the option market for the first time, there were also a number of agricultural water districts willing to bid up the price of water for application to other high-value crops. Higher contract prices may also result from a reduction in informational asymmetry. As information regarding the high cost of urban water shortage is disseminated, sellers adjust their price estimates. In short, the seller incorporation of new information regarding the high shortage cost of urban water is one possible explanation for the observed price trend.

### 3.8.5 Future Contracting

The viability of north-south water transfers and the social welfare gain from contracting depend critically on the marginal cost of electricity and on the expected spot market price for rice. The price of rice on the global exchange hit a record of \$24.10 in April of 2008. Past rice prices have been low compared to other crops cultivated in the state. If rice prices remain high, water transfers from other lower-value crops may become predominant. Commodity price increases are an important consideration in designing flexible future transfer contracts. An increase in the expected price of rice on the global exchange raises the seller's reserve value, at the same time lowering the total surplus from contracting. The contract price estimates for rice in 2008 ranged from \$13.5/af to \$19/af. At \$13.5/af, the marginal product of rice for water is \$84. At \$19/af, the marginal product of rice for water is \$203/af. The strike price for 2008 contracts is \$200/af, approximately equal to the opportunity cost of water to the seller. Therefore, at a base price of zero, a risk-neutral seller would prefer rice cultivation to a water sale. At a base price of \$50/af, however, the sale of water has a higher payoff than crop cultivation.

The system power rate has remained constant over the last three years and the State Water Project (SWP) charge has again been approved for \$110/af in 2009. Nonetheless, a future increase in the system power rates due to increasing statewide energy costs remains a distinct possibility. Table 1 reports the social welfare under contracting given variable expected prices for rice and three scenarios for system power costs.

## 3.9 Summary

At the beginning of this chapter we posed four questions: (1) what is the structure of the optimal option contracts, i.e., what are the quantities and prices associated with the

Price (\$/cwt)	MP of Water	Reserve	Case 1	Case 2	Case 3
0	0	0	18.19	8.08	6.79
5	0	0	18.19	8.08	6.79
10	7.65	0.77	17.43	7.31	6.02
15	115.98	11.60	6.59	0	0
20	224.31	22.43	0	0	0
25	332.65	33.26	0	0	0
30	440.98	44.10	0	0	0

Table 3.5: Sensitivities to the Price of Rice (in millions of dollars).

optimal contract? (2) how does the presence of uncertainty regarding the buyer's downstream demand and the price of the outside (market) option impact the structure of the contract? (3) how does the contract change when the buyer has an outside supply option, which could include access to a spot market, and (4) how does the allocation achieved under bilateral contracting compare to the socially optimal allocation?

In Sections 3.5-3.7 we derived the optimal contracts under three scenarios of interest – the first, where the buyer lacks an outside supply option, the second, where the buyer has an outside supply option available at a fixed price, and the third, where the buyer and seller have joint spot market access. The introduction of an outside option for the buyer results in lower contract prices and a generally more favorable outcome for the buyer, in keeping with our intuition. As we observed, this result holds *irrespective of whether the buyer actually expects to rely on the outside option once the contract is signed*.

Under joint spot market access, the seller prices the option more aggressively, reflecting the risk he faces in the second-period: if the market price realizes high, then the buyer will elect to call options rather than meet supply on the market. The seller would have preferred to receive the market price than the strike price, since  $m > s$ . Conversely, if the spot market price realizes low then the buyer will buy from the market instead of calling the options – and again the seller would have preferred the other outcome. With joint spot market access, the buyer's exercise decision (rather obviously) no longer independent from the market realization, as was the case when the buyer did not have market access.

The introduction of downstream demand uncertainty changes the structure of the optimal pricing rules (illustrated by comparison to the pricing rules derived in Wu et al. (2002)). We also found that it changes the conditions under which contracting is feasible. In Section 3.4, we established a sufficient condition for contracting with no buyer market access, namely that

the buyer's marginal utility of holding at least one option exceed the expected spot market price:  $u_Q(0, D) > \bar{m}$ . With joint spot market access and downstream demand uncertainty, we get feasible contracting whenever the buyer's potential loss from non-contracting is "large enough," where we observe  $(u_Q(Q, D))$  can be arbitrarily large.

With regard to the final question, we establish in Section 3.1 that bilateral contracting fails in general to achieve the first-best allocation. First-best can be achieved under an option contract with a zero upfront fee and a strike price indexed to the market price. Without indexing, even the contract established by a social-welfare maximizing board, with prices  $(0, \bar{m})$ , is inefficient. The seller-set contract, which generally levies an upfront fee, further distorts the optimal allocation.

Is the allocative inefficiency of option contracting a concern? The answer from an economic standpoint must be yes. Whether it is our chief concern is a more nuanced question. As an instrument facilitating trade, we must conclude that option contracting is still desirable when trade is desirable (and when the alternative is no trade). In the California water market, where trade has been stalled, option contracting offers a politically viable way to stimulate transactions and move us closer to an efficient reallocation of water resources

We close by observing that our focus on allocative efficiency suggests its primacy over distributional objectives, which may not be the case. In fact, it is bound *not* to be the case in political or social spheres. As discussed, the surplus-maximizing contract is the one that sets a zero upfront fee and then indexes the strike to the true market price. This contract transfers all of the gains from trade (surplus) to the buyer. While this is "efficient," we can hardly claim that it meets distributional objectives. This contract would never voluntarily be set by the seller (the farmer), since it amounts to voluntary forfeiture of a property right (or water right). If implemented by a legislative body, it would undermine current legal doctrine protecting individual property rights, not to mention make for a highly unpopular course of political action.

However, this contract could be advantageous to both parties when combined with a two-part tariff. This opens the door to a discussion of more sophisticated contracting designs, namely contracts in which options are fully indexed and fixed transfers are levied to appropriate (some portion of) the buyer's total surplus.

### 3.10 Appendix

**Proof of Proposition 3.2.** The sum of the buyer's and seller's expected payoffs is given as follows:

$$U(Q, s) + (s - \bar{m})y + \bar{m}K$$

where the transfer  $pQ$  cancels and  $y(z, s) = \int_{D(s)}^{\bar{D}(s, Q)} z(s, D)dF(D) + \int_{\bar{D}(s, Q)}^{\kappa} QdF(D)$ . Taking the FOC w.r.t.  $s$  we have

$$\begin{aligned} & \int_{D(s)}^{\bar{D}(s, Q)} [u_z(z(s, D), D)z_s - sz_s - z]dF(D) + \int_{\bar{D}(s, Q)}^{\kappa} [u_Q(Q, D)Q_s - sQ_s - Q]dF(D) \\ & + y + (s - \bar{m}) \left[ \int_{D(s)}^{\kappa} z_s(s, D)dF(D) + Q_s(1 - F(\bar{D}(s, Q))) \right] = 0 \end{aligned}$$

since  $y_s = \int_{D(s)}^{\kappa} z_s(s, D)dF(D)$ . A number of terms cancel and we are left with

$$\int_{D(s)}^{\bar{D}(s, Q)} z_s[u_z - \bar{m}]dF(D) + \int_{\bar{D}(s, Q)}^{\kappa} z_s[u_Q - \bar{m}]dF(D) = 0$$

This implies

$$u_z(z, D) = u_Q(Q, D) = \bar{m}$$

The joint surplus is maximized when the buyer's marginal value with respect to his exercise decision  $z$  is equal to the expected spot market price  $\bar{m}$  and the buyer's marginal value at  $Q$  is equal to the expected spot market price  $\bar{m}$ . Since  $u(Q, s)$  increasing in  $Q$ , setting  $p^{**} = 0$  guarantees  $Q = z(s, \beta)$ , where  $\beta$  is the upper bound of the support of demand  $\bar{D}$ . That is, at an option price of zero the buyer holds as many options as he would ever exercise. Hence, for  $p^{**} = 0$ , the condition  $u_z(z, D) = u_Q(Q, D)$  holds.

To find the  $s$  that sets the buyer's marginal value equal to the expected spot market price, we recall that the buyer's second-period exercise rule  $z$  is given as the solution to the following maximization problem:

$$\begin{aligned} & \max_z \{u(z, D) - sz\} \\ & \text{s.t. } z \leq Q \end{aligned}$$

Hence

$$z(s, D) = \begin{cases} a(s; D) & \text{if } a(s; D) \leq Q \\ Q & \text{otherwise} \end{cases}$$

where  $a(x; D) = \{z \in \mathbb{R}_+ : u_z(z, D) = x\}$ . At  $p^{**} = 0$  we have  $a(s; D) \leq Q \forall D$ . Hence surplus is maximized by setting  $s^{**} = \bar{m}$ . ■

**Proof of Proposition 3.2.** The proof proceeds in two parts:

*Part 1.* First we establish that the seller contracts at any  $s > \bar{m}$ , with  $p = 0$ . The seller will contract at any price pair  $(p, s)$  that results in a higher payoff than her reserve value  $\bar{m}K$ . Let  $\hat{s} = \bar{m} + \epsilon$  for some  $\epsilon > 0$  and let  $y$  denote the expected number of options to be exercised at  $\hat{s}$ . Then the seller's payoff at  $(0, \hat{s})$  is given as  $(\bar{m} + \epsilon)y + \bar{m}(K - y) = \bar{m}K + \epsilon y > \bar{m}K$ .

*Part 2.* Now we establish willingness to contract on the buyer's part at  $(0, \bar{m} + \epsilon)$ , for some  $\epsilon > 0$ . The buyer's payoff from contracting at some  $s$ , with  $p = 0$ , is given as

$$\int_0^{D(s)} u(0, D) dF(D) + \int_{D(s)}^{\bar{D}(s, Q)} u(g(s, D), D) - sg(s, D) dF(D) + \int_{\bar{D}(s, Q)}^{\kappa} u(Q, D) - sQ \quad (3.31)$$

The buyer's reserve value is given as

$$\int_0^{\kappa} u(0, D) dF(D). \quad (3.32)$$

Then, subtracting (3.32) from (3.31) we get the buyer's necessary condition for contracting:

$$\int_{D(s)}^{\bar{D}(s, Q)} [u(g(s, D), D) - sg(s, D) - u(0, D)] dF(D) + \int_{\bar{D}(s, Q)}^{\kappa} [u(Q, D) - sQ - u(0, D)] dF(D) \geq 0. \quad (3.33)$$

The integrand in both integrals is by definition non-negative, since the buyer will never elect to exercise an option for which the marginal value is below . In order for the strict equality to hold in 3.33 there must exist a  $D$  satisfying  $\bar{D} = \{D : g(s, D) = Q\}$ . Otherwise the two integrals in the expression 3.33 collapse. Replacing  $s$  with  $\bar{m}$  we find that there exists a demand level  $D$  for which the buyer is willing to pay at least  $\bar{m}$  to execute the option. ■

**Proof of Lemma 3.3.** We have:

$$Q_p = \frac{h'(\cdot)}{1 + h'(\cdot)sf(\bar{D})\bar{D}_Q}$$

$$Q_s = \frac{h'(\cdot)(1 - F(\bar{D}))}{1 + h'(\cdot)sf(\bar{D})\bar{D}_Q}$$

where  $h'(\cdot) = 1/[-\bar{D}_Q[V'(\bar{D}, D)f(\bar{D})] + \int_{\bar{D}(s, Q)}^{\infty} V''(Q, D)f(D)dD] < 0$ . We observe that  $V'(\bar{D}, D) \geq s$ , which implies  $|h'(\cdot)sf(\bar{D})\bar{D}_Q| \leq 1$  and, hence, the denominators of  $Q_p$  and  $Q_s$  are positive. Hence,  $Q_p, Q_s \leq 0$ , since  $h'(\cdot) < 0$ . ■

**Proof of Lemma 3.4.** Direct from Lemma 3.3. ■

**Proof of Lemma 3.6.** Let  $u_z(0, D) = \omega < \infty$ . Assume  $s > \omega$ . Then, we have  $u_z \leq \omega = s > \omega$ , which is a contradiction. (The first equality comes from the concavity of  $u$  ( $u_z$  decreasing) and the assumption  $u_z(0, D) = \omega < \infty$ , and  $u_z = s$  comes from the buyer's FOC.) Hence,  $s \leq \omega$ . By similar reasoning, we must have  $p \leq \omega$ , since the *most* the buyer is willing to pay (in total) for a unit of supply is  $u_z \leq \omega$ . Hence  $p + s \leq \omega$  and, by Lemma 3.5 above,  $s > 0$  implies  $p < \omega$ . ■

**Proof of Theorem 3.5.** We are maximizing a continuous function on a compact set, where compactness comes from Lemmas 3.5 and 3.6, which ensure that  $p \in [0, \gamma]$  and  $s \in [0, \gamma]$ . We are therefore guaranteed existence of a maximizer by Wierstrass. The uniqueness of these critical points comes from (3.10) and (3.11) – the two equations providing the necessary conditions which have unique solutions, namely  $s^* = m - \frac{\bar{g}}{\int_D^{\bar{D}} g_s f(D) dD}$  and  $p^* = \frac{-(s-m)(1-F)}{1-1/\varepsilon_p}$ . Assumptions A1 and A2 assure us joint concavity of  $V$  in  $(p, s)$  and, therefore, that the critical points are global maximizers.

The derivations of  $s^*$  and  $p^*$  are given below:

*Step 1.* Substituting our expressions for  $y_p$  and  $y_s$ , given in (3.12) and (3.13) above, respectively, into the seller's FOCs given in (3.10) and (3.11), we have

$$V_p = [p + (1 - F(\bar{D}))(s - m)]Q_p + Q = 0$$

$$V_s = [p + (1 - F(\bar{D}))(s - m)]Q_s + (s - m) \left[ \int_D^{\bar{D}} g_s f(D) dD \right] + y = 0$$

*Step 2.* The first equation yields an expression for  $p$ ,  $p = \frac{-Q}{Q_p} - (s - m)(1 - F(\bar{D}))$ , which we substitute into the expression for  $V_s$ :

$$V_s = -Q \frac{Q_s}{Q_p} + (s - m) \left[ \int_D^{\bar{D}} g_s f(D) dD \right] + y = 0$$

This implies

$$Q = \left( y + (s - m) \left[ \int_D^{\bar{D}} g_s f(D) dD \right] \right) \frac{Q_p}{Q_s}$$

*Step 3.* From Lemma 3.4, we have  $\frac{Q_p}{Q_s} = \frac{1}{1-F(\bar{D})}$ . Using this, and our definition of  $y$ ,  $y = Q(1 - F(\bar{D})) + \bar{g}$ , our expression above for  $V_s$  becomes

$$Q = \left[ Q(1 - F(\bar{D})) + \bar{g} + (s - m) \left( \int_D^{\bar{D}} g_s f(D) dD \right) \right] \frac{1}{1 - F(\bar{D})}$$

This implies

$$-(s - m) \left[ \int_D^{\bar{D}} g_s f(D) dD \right] = \bar{g}$$

This implies

$$s - m = -\frac{\bar{g}}{\int_D^{\bar{D}} g_s f(D) dD}$$

which implies

$$s = m - \frac{\bar{g}}{\int_D^{\bar{D}} g_s f(D) dD} \geq m$$

since  $\bar{g} \geq 0$  and  $\int g_s \leq 0$ .

*Step 4.* To derive  $p^*$  we revisit the expression above for  $p$ :  $p = -\frac{Q}{\varepsilon_p} - (s - m)(1 - F(\bar{D}))$ . We define the (option) price elasticity of demand  $Q$  as  $\varepsilon_p = -\frac{\partial Q}{\partial p} \cdot \frac{p}{Q}$ . Then, by substitution,

$$p^* = 1/\varepsilon_p \cdot p - (s - m)(1 - F)$$

implies

$$p = \frac{-(s - m)(1 - F)}{1 - 1/\varepsilon_p}$$

■

**Proof of 1.** The proof follows directly from Theorem 3.5 and Lemma 3.5. From Theorem 3.5, we have  $s \geq m$ , which implies the numerator in the expression for  $p^*$  is negative. Then, by Lemma 3.5,  $p > 0$ , so we must have the denominator in the expression for  $p$  also negative:  $1 - 1/\varepsilon_p \leq 0$  implies  $\varepsilon_p \leq 1$ . ■

**Proof of Theorem 3.7.** The result follows directly from the buyer's FOC with respect to  $Q$ , where the buyer's payoff function is given in (3.21). ■

**Proof of Theorem 3.6.** The proof follows the same lines as the proof of Theorem 3.5, where our expressions for  $V_p$  and  $V_s$ , the seller's FOCs, are now given as

$$V_p = [p + (1 - F(\bar{D}))(s - m)]Q_p + Q - \lambda\phi_p = 0$$

$$V_s = [p + (1 - F(\bar{D}))(s - m)]Q_s + (s - m) \left[ \int_D^{\bar{D}} g_s f(D) dD \right] + y - \lambda\phi_s = 0$$

where  $\lambda$  is the Lagrange multiplier on the buyer's IR constraint. ■



**Proof of Lemma 3.9.** Let  $(\hat{p}^*, \hat{s}^*)$  denote the prices charged by the seller to the buyer *with* an outside option, and let  $(p^*, s^*)$  denote the optimal prices charged *without* an outside option. Also,  $\hat{Q}^*$  denotes demand with an outside option and  $Q^*$  denotes demand without an outside option. In the case where the buyer has a outside option, the optimal price pair charged by the seller is (weakly) smaller, and since  $Q$  is strictly decreasing in  $p$  and  $s$ , a weakly smaller  $(p, s)$  implies a higher  $Q$ . We have two cases to consider: the first where the buyer's IR constraint is non-binding and the second where it is in fact binding.

*Part 1. Non-binding IR Constraint*

By Theorem 3.6, the seller applies the same general price-setting rules both when the buyer has an outside option and when he does not (when the IR constraint is non-binding). The seller does not necessarily set the same prices, however, since he faces a different demand function. As established in Lemma 3.7, the demand function he faces is piecewise. On the first piece of the demand curve, the seller faces the *same demand* as the case with no outside option – however pricing on this piece of the demand curve is now subject to the constraint  $\bar{D}(p_a, Q(p, s)) > \kappa$ , which implies lower prices (since  $Q$  strictly decreasing in  $p$  and  $s$  and  $\bar{D}$  strictly increasing in  $Q$ ). On the second piece of the demand curve, demand is stronger (recall  $h$  a decreasing function and hence  $\hat{Q} = [\hat{h}(p + s(1 - F(\bar{D}(s, Q)))) - p_a(1 - F(\bar{D}(p_a, Q)))]_+ > [h(p + s(1 - F(\bar{D}(s, Q))))]_+ = Q$ ). Suppose  $\hat{Q}^* < Q^*$ . There are then higher prices  $(p, s)$  at which the seller could have achieved  $Q^* = \hat{Q}^*$  and been strictly better off, which contradicts the optimality of  $(p^*, s^*)$ . Hence we must have  $\hat{Q}^* \geq Q^*$ .

*Part 2. Binding IR Constraint*

Bindingness of the buyer's IR constraint implies lower prices: the constraint is binding in  $Q$ , and  $Q$  is strictly decreasing in  $p$  and  $s$ . (See Theorem 3.6.) To see that prices must also be weakly lower than those in the case without an outside option, we observe that from (3.25) the buyer's reserve is greater in the presence of an outside option and, hence, the IR constraint is stricter. ■

**Proof of Proposition 3.3.** The proof proceeds in four parts.

*Part 1.* The buyer's FOC is given as

$$\begin{aligned} P(m \geq s - \tau) & \left[ -\frac{\partial \bar{D}(s, Q)}{\partial Q} [u(g(s, \bar{D}), \bar{D}) - sg(s, \bar{D})] dF(D) \right] \\ & - \int_{s-\tau}^{\beta} \frac{\partial \bar{D}(s, Q)}{\partial Q} [u(Q, \bar{D}) - sQ] dF(D) dF(m) \\ & \int_{s-\tau}^{\beta} \frac{\partial \bar{D}(m + \tau, Q)}{\partial Q} [u(Q, \bar{D}) - sQ] dF(D) dF(m) \end{aligned}$$

$$\begin{aligned}
& - \int_{s-\tau}^{\beta} \frac{\partial \bar{D}(m+\tau, Q)}{\partial Q} [u(g(m+\tau, \bar{D}), \bar{D}) - sg(m+\tau, \bar{D})] dF(D) dF(m) \\
& + \int_{s-\tau}^{\beta} \int_{\bar{D}(s, Q)}^{\bar{D}(m+\tau, Q)} [u_Q(Q, D) - s] dF(D) dF(m) + \int_{s-\tau}^{\beta} \int_{\bar{D}(m+\tau, Q)}^{\kappa} [-s + (m+\tau)] dF(D) dF(m) - p
\end{aligned}$$

The first four terms cancel, and we are left with

$$\begin{aligned}
& \int_{s-\tau}^{\beta} \int_{\bar{D}(s, Q)}^{\bar{D}(m+\tau, Q)} [u_Q(Q, D)] f_D(D) f_m(m) dm dD = \\
& p + sP(\bar{D}(s, Q) \leq D < \bar{D}(m+\tau, Q) | m \geq s-\tau) + (s - (m+\tau))P(D \geq \bar{D}(m+\tau, Q) | m \geq s-\tau).
\end{aligned} \tag{3.34}$$

*Part 2.* We simplify the expression for

$$P(\bar{D}(s, Q) \leq D < \bar{D}(m+\tau, Q) | m \geq s-\tau) \tag{3.35}$$

in terms of the joint cdf of  $\tilde{m}$  and  $\tilde{D}$ , which we denote  $\hat{F}$  and the cdf of  $\tilde{m}$ , which we denote  $\hat{F}_m$ .

Recall that Bayes' Rule establishes the following relationship for conditional probabilities:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} = \frac{P(A \cap B)}{P(B)}$$

By application of Bayes' Rule, we can rewrite (3.35) as

$$\frac{P(\bar{D}(s, Q) \leq D < \bar{D}(m+\tau, Q) \ \& \ m \geq s-\tau)}{P(m \geq s-\tau)}$$

where we can now rewrite the numerator and denominator in terms of the joint cdf of  $\tilde{D}$  and  $\tilde{m}$  and the cdf of  $\tilde{m}$ , respectively, as follows:

$$\frac{(\hat{F}(\gamma, \bar{D}(m+\tau, Q)) - \hat{F}(s-\tau, \bar{D}(m+\tau, Q))) - (\hat{F}(\gamma, \bar{D}(s, Q)) - \hat{F}(s-\tau, \bar{D}(s, Q)))}{1 - \hat{F}_m(s-\tau)}$$

where  $\gamma$  is the upper bound of the support of the market price  $\tilde{m}$ . Letting  $C = \hat{F}(\gamma, \bar{D}(s, Q)) - \hat{F}(s-\tau, \bar{D}(s, Q))$  and  $B = \hat{F}(\gamma, \bar{D}(m+\tau, Q)) - \hat{F}(s-\tau, \bar{D}(m+\tau, Q))$ , we see that (3.35) can be written more succinctly as

$$\frac{B - C}{1 - \hat{F}_m} \tag{3.36}$$

*Part 2.* By a second application of Bayes' Rule, we can simplify the expression for

$$P(D \geq \bar{D}(m+\tau, Q) | m \geq s-\tau). \tag{3.37}$$

We arrive at the following expression

$$\frac{1 - B}{1 - \hat{F}_m(s - \tau)}, \quad (3.38)$$

where  $B$  is as defined above

*Part 3.* Substituting our simplified expressions for the two probabilities into (3.34), we can then express the optimal  $Q$  as

$$Q^* = \left[ h \left( p + s \left[ \frac{B - C}{1 - \hat{F}_m} \right] + (s - (m + \tau)) \left[ \frac{1 - B}{1 - \hat{F}_m(s - \tau)} \right] \right) \right]_+$$

which simplifies to

$$Q = \left[ h \left( p + s \left[ \frac{1 - C}{1 - \hat{F}_m} \right] - (m + \tau) \left[ \frac{1 - B}{1 - \hat{F}_m(s - \tau)} \right] \right) \right]_+$$

where  $h(x) = \{Q \in \mathbb{R}_+ : \bar{u}(Q; s, D) = x\}$  and  $\bar{u}(s, Q, D) = \int_{s-\tau}^{\beta} \int_{\bar{D}(s, Q)}^{\bar{D}(m+\tau, Q)} [u_Q(Q, D)] dF(D) dF(m)$  and  $C = \hat{F}(\gamma, \bar{D}(s, Q)) - \hat{F}(s - \tau, \bar{D}(s, Q))$ ,  $B = \hat{F}(\gamma, \bar{D}(m + \tau, Q)) - \hat{F}(s - \tau, \bar{D}(m + \tau, Q))$ , and  $\hat{F}$  is the joint cdf of  $\tilde{m}$  and  $\tilde{D}$  and  $\hat{F}_m$  is the cdf of  $m$  and  $\gamma$  is the upper bound of the market price  $\tilde{m}$ . ■

**Proof of Theorem 3.8.** We execute the proof in three parts:

*Part 1.* From (3.30) the seller's objective function is given as

$$V = pQ + P(m \geq s - \tau)(s - E[m|m \geq s - \tau])y + (\bar{m} - b)K$$

where  $\hat{m} = E[m|m \geq s - \tau]$ .

Then the seller's FOCs are given as

$$V_p = pQ_p + Q + P(m \geq s - \tau)(s - \hat{m})y_p = 0 \quad (3.39)$$

$$\begin{aligned} V_s = pQ_s + [P(m \geq s - \tau) + (s - \tau)f(s - \tau)]y + P(m \geq s - \tau)(s - \hat{m})y_s \\ + \left[ \frac{\partial}{\partial s}(P(m \geq s - \tau)) \right] sy = 0, \end{aligned} \quad (3.40)$$

where the derivative of the conditional expectation  $\hat{m}P(m \geq s - \tau)$  of the spot market price is given as:

$$\frac{\partial}{\partial s} \left[ \int_{s-\tau}^{\beta} mf(m)dm \right] = [\tau - s]f(s - \tau).$$

*Part 2.* Then substituting in  $y_s = Q_s(1 - F(\bar{D}(s, Q))) + \int_{D(s)}^{\bar{D}(s, Q)} g_s(s, D)dF(D)$  we have

$$V_s = Q_s[p + (1 - F)P(m \geq s - \tau)(s - \hat{m})] + P(m \geq s - \tau)(s - \hat{m})\hat{g}_s + y[sP_s + P(m \geq s - \tau) + (s - \tau)f(s - \tau)] = 0$$

where  $P_s = \frac{\partial}{\partial s}(P(m \geq s - \tau))$  and  $\hat{g}_s = \int_{D(s)}^{\bar{D}(s, Q)} g_s dF(D)$ . Then from (3.39) (with  $Q_p(1 - F)$  substituted for  $y_p$ ) we have an expression for  $p$ :  $p = -\frac{Q}{Q_p} - (1 - F)P(m \geq s - \tau)(s - \hat{m})$ .

Hence,

$$\begin{aligned} V_s &= Q_s \left( -\frac{Q}{Q_p} - (1 - F)P(m \geq s - \tau)(s - \hat{m}) + (1 - F)P(m \geq s - \tau)(s - \hat{m}) \right) \\ &\quad + P(m \geq s - \tau)(s - \hat{m})\hat{g}_s + y[sP_s + P(m \geq s - \tau) + (s - \tau)f(s - \tau)] \\ &= -Q\frac{Q_s}{Q_p} + P(m \geq s - \tau)(s - \hat{m})\hat{g}_s + y[sP_s + P(m \geq s - \tau) + (s - \tau)f(s - \tau)] = 0 \end{aligned}$$

which we re-write as

$$-Q\Upsilon + P(m \geq s - \tau)(s - \hat{m})\hat{g}_s + y[s(P_s + f(s - \tau)) + \Gamma] = 0$$

where  $\Gamma = (P(m \geq s - \tau) - \tau f(s - \tau))$  and  $\Upsilon = \frac{Q_s}{Q_p}$ . Then we have

$$s[(P_s - f(s - \tau))y + P(m \geq s - \tau)\hat{g}_s] = \hat{m}P(m \geq s - \tau)\hat{g}_s - y\Gamma + Q\Upsilon$$

from which we arrive at our expression for  $s^*$ :

$$s^* = \frac{\hat{m}P(m \geq s - \tau)\hat{g}_s - y\Gamma + Q\Upsilon}{(P_s + f(s - \tau))y + P(m \geq s - \tau)\hat{g}_s}$$

NOTE: For  $\tau = 0$ , the optimal  $s$  is given as

$$s^* = \frac{P(m \geq s)(\hat{m}\hat{g}_s - y) + Q(1 - F)}{(P_s + f(s))y + P(m \geq s)\hat{g}_s}$$

where for  $(P_s + f(s)) \ll 1$  the optimal  $s^*$  is *approximately* given as

$$s^* = \hat{m} + \frac{Q(\Upsilon - P(m \geq s)(1 - F)) - \bar{g}P(m \geq s)}{P(m \geq s)\hat{g}_s}$$

where the denominator is negative (since  $\hat{g}_s < 0$ ) and whether the seller prices above or below the conditional expectation of  $\tilde{m}$  depends on how likely it is that all  $Q$  options are exercised (given by the probability  $(1 - F)$  and how likely it is that the market price realizes higher than  $s$ , where clearly  $Q > \bar{g}$ ).

*Part 3.* For the case where the seller's FOCs are necessary and sufficient, we can derive the optimal  $p$  from our expression  $p = -\frac{Q}{Q_p} - (s - \hat{m})P(m \geq s - \tau)(1 - F(\bar{D}(s, Q)))$ . We define

the (option) price elasticity of demand  $Q$  as  $\varepsilon_p = -\frac{\partial Q}{\partial p} \cdot \frac{p}{Q}$ . Then  $p^* = 1/\varepsilon_p p - (s - \hat{m})P(m \geq s - \tau)(1 - F(\bar{D}(s, Q)))p^* = \frac{-(s - \hat{m})(1 - F)P(m \geq s - \tau)}{1 - 1/\varepsilon_p}$ .

■

# Chapter 4

## Option Contract Competition

### 4.1 Motivation

While bilateral contracting has been the standard in the water market to date, it is reasonable to assume that as the market develops, multilateral contracting will become more commonplace. The possibility of multiple buyers and sellers active in the marketplace suggests the need to consider option contract competition.<sup>1</sup> In particular, it motivates our consideration of two extensions to the standard bilateral option contracting model developed in Chapter 3. We consider the following two scenarios: a single seller facing multiple buyers and a single buyer facing multiple sellers. The case of multilateral contracts between multiple sellers *and* buyers is left for future research. In general, we are interested in gaining insight into shifts in prices and quantities given the presence of additional buyers and sellers, and the corresponding implications for surplus distribution. For example, rather intuitively the presence of additional sellers leads to price competition that is favorable to the buyer.

A single seller facing multiple buyers will continue to invoke the pricing rules established in Chapter 3 for seller-set bilateral contracts. For the purpose of deriving the seller's pricing decision, aggregate demand simply replaces the ordinary demand function in the model. Individual demand is now served in order of willingness-to-pay. The presence of additional buyers implies the possibility that some buyers (with lower willingness to pay) will not get served at all. The pricing rules remain the same. However, given that demand is stronger (aggregate demand greater than individual demand), prices are higher. The buyer

---

<sup>1</sup>Buyer-seller matching will ultimately depend on infrastructure access by parties wishing to transact. In this sense, the market is likely to remain incomplete.

competition hence leads to higher prices and lower reliability for the average buyer. In the water market, this scenario corresponds to cases where a single irrigation district (a seller) faces several potential buyers, possibly including urban water districts or other agricultural districts. The dependence on infrastructure access increases the likelihood of this scenario, limiting the option for a given set of buyers such that they are likely to want to purchase water from the same single seller.

When there are multiple sellers, the contracting model needs to account for the strategic interaction between them. Bertrand (1883) first introduced a model of price competition wherein sellers (or firms) with constant returns to scale technologies simultaneously set prices. Under the assumption of symmetric costs, the firms price competitively (price equal to marginal cost) and obtain zero profit. The result, which relies on an undercutting argument in which one player offering a price pair above marginal cost leads the other player to adjust his price by  $\varepsilon$  and thereby win the whole market until prices are driven to marginal cost and profits are driven to zero, is not robust. The introduction of asymmetries in cost, product differentiation, and capacity constraints leads to prices above marginal cost and positive profits.

A number of variations of the standard Bertrand model have been considered, including asymmetric costs and product differentiation between firms, as well as the alternative strategic decision of quantity instead of price. The latter is referred to as Cournot competition, in reference to a model by Cournot (1838) in which firms choose a production level and prices are then assumed to be set such that the market clears. A third variation, of particular interest to us, is the presence of capacity constraints. In reality, sellers have fixed (finite) capacity in the short-run. We treat the case of capacity-constrained Bertrand competition in a duopoly, first under complete information and then under incomplete information. Both cases are applicable to the water market. In the first (complete information), the sellers know each others capacities. Water rights and water deliveries are public information, making the assumption of complete information appropriate in some cases. However, irrigation districts negotiating trades will not offload their entire capacity, or supply, as they will keep a percentage of land in production. The actual amount that a district will unload is uncertain, and the assumption of incomplete information applies in this scenario.

#### 4.1.1 Related Literature

Osborne and Pitchik (1986), henceforth referred to as O&P, treat the case of capacity-constrained Bertrand competition in a duopoly. They find that pure-strategy equilibrium

exist when there is a state of “over-capacity” or “under-capacity” in the supply system relative to demand. That is, when capacities are either small (under-capacity) or large (over-capacity) the model has an equilibrium in pure strategies. When total capacity is in a mid-range, there are equilibria in mixed strategies. In the state of under-capacity, the total supply available from the two sellers is *less than* that which would be demanded at the monopoly price. The two sellers then set a price  $p$  such that the total capacity clears, i.e.,  $p = 1 - k_1 - k_2$ , where  $k_1$  and  $k_2$  are the individual capacities of seller one and two, respectively. Neither seller has an incentive to deviate from this price, since lowering price would result in the same volume of sales (total capacity) but lower profit. Raising price would result in a sub-optimal decrease in the amount sold (below capacity) given that at the monopoly price, or profit-maximizing price, total system capacity is offloaded. In the state of over-capacity, the seller with the larger capacity can fully meet demand at the monopoly price. At any price above marginal cost, the other seller has an incentive to deviate, undercutting price to serve the total market. The pure-strategy equilibrium in the over-capacity state is that in which both players charge marginal cost and profits are zero. In general, equilibrium prices and capacity move in opposite directions, with smaller capacity translating into higher prices.

The O&P analysis assumes that demand is “efficiently rationed,” where the low-price firm serves demand up to capacity first. The residual demand served by the high-price firm  $i$  is given as  $\min\{\max\{0, Q(p_i) - k_j\}, k_i\}$  where  $k_i$  is the capacity of firm  $i$  and, likewise,  $k_j$  is the capacity of firm  $j$ , and  $Q(p_i)$  is market demand at price  $p_i$ . The choice of the rationing rule is important to the analysis. Allen and Hedwig (1993) treat the case of proportional rationing, in which there is some probability that the low-price firm is unable to meet demand. In this case, residual demand served by the higher-price firm is given as  $\max\{0, \min\{Q(p_i)(1 - k_j/Q(p_j)), k_i\}\}$ , where  $(1 - k_j/Q(p_j))$  is the probability that a buyer cannot fill demand from seller  $j$ . They demonstrate that there exists a Nash equilibrium in nondegenerate mixed strategies. A nondegenerate mixed strategy is one in which the distribution over possible actions that constitutes a strategy does not contain any atoms, e.g., the distribution is smooth.

Another extension of the Bertrand model, by Cayseele and Furth (1996) introduces the possibility of buyouts or first-refusal contracts. Under a buyout, the high-price firm would purchase the entire capacity of the low-price firm and then proceed to meet all of the market demand (assuming demand is less than capacity) at the high-price. The first-refusal contract is structured in a similar fashion, with the high-price firm purchasing the total *demand* from the low-price firm (as opposed to the total *capacity* of the low-price firm). In keeping with



O&P, Cayseele and Furth find that the existence of pure strategy equilibrium depends on the degree of under-capacity. Specifically, only if the total joint capacity on the market is less than the monopoly output is there a pure strategy equilibrium.

The O&P analysis is preceded by the work of Kreps and Scheinkman (1983), which considers the joint strategic decisions of quantity and price in a two-stage setting, with quantity set in the first stage, followed by price in the second stage. Kreps and Scheinkman show that under the assumption of concave demand and efficient rationing, the equilibrium involves production levels and capacities that correspond to the Cournot output levels and market-clearing prices. In other words, the Cournot result holds. The results is not robust to a change in the assumed rationing rule, e.g., under proportional versus efficient rationing the result no longer holds. Benoit and Krishna (1987) consider a dynamic model of duopoly in which firms set both prices and quantities. In this dynamic setting, they find that when capacity is inflexible firms tend to carry excess capacity and profits are below monopoly levels. If capacity is flexible, joint profits at monopoly levels can be obtained.

The Bertrand model has also been extended to account for uncertainty by both Reynolds and Wilson (2000) and Janssen and Rasmusen (2002). Reynolds and Wilson (2000) treat the case of uncertain demand. In their two-stage model, firms first (in stage one) choose their production levels while demand is unknown and then (in stage two) set their prices once demand is known. For demand variability exceeding a given threshold, Reynolds and Wilson establish the nonexistence of a pure-strategy equilibrium in capacities. With regard to prices, they show that output prices have a positive variance for low demand and zero variance for high demand. In other words, when demand realizes high and the firms can fully offload their joint capacities, there exist pure strategy equilibria (as was the case in O&P). If the demand realization is low, the equilibrium may be in mixed strategies, with total demand less than capacity. Janssen and Rasmusen (2002) introduce uncertainty in a different form. They treat the case of Bertrand competition where there is a positive probability that a given firm is inactive. This model captures the possibility that a given firm is saturated in orders and thus not an active competitor at some point in time. The example Janssen and Rasmusen provide is of a carpenter asked by a homeowner to submit a bid for a renovation project. The carpenter assumes that the homeowner will solicit other bids. However, he also assigns a positive probability to the “inactivity” of these other carpenters, e.g., he assigns a positive probability to possibility that they are fully booked for the month. Janssen and Rasmusen establish the existence of mixed-strategy equilibrium in the model, with total industry profits declining in the number of firms.

To the best of our knowledge, there has been no treatment to date of capacity-constrained

Bertrand competition in a setting of incomplete information regarding capacities. In Section 4.3.2 we develop a two-player model to treat this scenario. The capacity of the players is assumed to be either “high” or “low,” e.g., the capacity of player  $i$  is  $k_i \in \{k_H, k_L\}$ . As we show, unlike the case with complete information where the two players charge the same price in the symmetric pure-strategy equilibrium, here the players charge different prices based on their capacity levels. The price charged by the firm in the high-capacity state is lower than the price charged by the low-capacity firm. At the pure-strategy equilibrium, joint capacity is fully offloaded in *both* the complete information and the incomplete information case.

Analogously, we can identify regions where there is under-capacity or over-capacity in the water contracting market and pure-strategy equilibria associated with these two states for the two-seller case. In a state of under-capacity, the model predicts that sellers set prices such that total capacity is demanded (and hence there is no incentive for price under-cutting). In the over-capacity state, the sellers will charge their reserve values, making zero profits. Both of these predicted outcomes have intuitive appeal in so far as we expect rents to be increasing in market power, where market power can be characterized in terms of capacity relative to demand. (The lower the ratio, the greater the collective market power and the higher the rents.) In the mid-range capacity states, the model predicts mixed strategy equilibria. In general, these equilibria lack intuitive appeal, as they imply that a given seller assigns a probability to each possible action (here, price) as his strategy.

As to what actually happens in these mid-ranges, we can posit several possibilities. Perhaps the two sellers act independently as follows: each considers his profit-maximizing price (as if he were the only seller) and then, recognizing the presence of competition, he shaves his price by some amount (an amount that seems reasonable to him). Then whoever “shaved” the most will meet the bulk (if not all) of demand, and it is impossible to predict ahead of time which might be the seller who sets the lower price. Alternatively, there may be a collusive outcome, where the sellers agree to set a price that provides what they consider a “minimally acceptable” level of utility to the buyer, where the “minimally acceptable” utility level is high enough so as to avoid political recourse (and clearly above the seller’s reserve). A third possibility is a three-way negotiation initiated by one of the two sellers. There are numerous other possibilities, highlighting the difficulty in predicting an outcome without additional information regarding the sellers’ relationship (friendly vs. antagonistic vs. anonymous), the legal statutes in place, precedents of past behavior, and other salient factors.

### 4.1.2 Outline

The rest of this chapter is organized as follows. In Section 4.2 we discuss the case where a single seller faces multiple buyers. For example, an irrigation district in the California water market may have the opportunity to offload water to several willing buyers, including nearby irrigation districts as well as urban water districts. Then in Section 4.3 we treat the case with multiple sellers, discussing first the case with two sellers and complete information. We assume that demand is rationed efficiently. We consider the case with complete information and show that under a simplified demand structure, e.g., demand equal to zero or one, the two-part price structure  $(p, s)$  reduces to a single tariff  $r = p + \theta s$ , where  $\theta$  is the probability that demand is equal to one. The analysis from O&P for capacity-constrained Bertrand competition then applies. In particular, there exist pure strategy equilibria in the state of under-capacity and over-capacity, as discussed. In both cases, the two players levy the same tariff in equilibrium. In the case with under-capacity, the players make a profit; in the case of over-capacity, the players set prices equal to marginal cost and break-even. We then consider the case with incomplete information for two players. Player  $i$  is either of type “high-capacity” or “low-capacity.” From player  $j$ ’s perspective, both capacity types are equally likely and are independently distributed from  $j$ ’s type. We show that in this case, there exists a pure-strategy equilibrium, where the player in the high-capacity state ( $k_H$ ) charges a price below that charged by the player in the low-capacity state ( $k_L$ ). In other words, the equilibrium prices are a decreasing function of the player’s capacity. Furthermore, the prices charged in the incomplete information case are below those charged in the complete information case. We close Section 4.3 with a brief discussion of an extension to a multi-seller ( $N > 2$ ) setting, modelled as a capacity-dependent auction. We conclude the chapter discussion in Section 4.4.

## 4.2 Multiple Buyers

The introduction of multiple buyers does not fundamentally change the option contracting model. As Wu et al. (2002) discuss, replacing the individual demand function with the aggregate demand function, the analysis of the seller-optimal pricing rules proceeds as in Chapter 3. Figure 4.1 depicts the aggregate demand curves in the case of homogeneous and heterogeneous players. As before, the seller selects an optimal  $(p, s)$  to offer, given the aggregate demand function  $\bar{Q}(p, s)$ . Figure 4.1 illustrates the demand function for homogeneous and heterogeneous players.

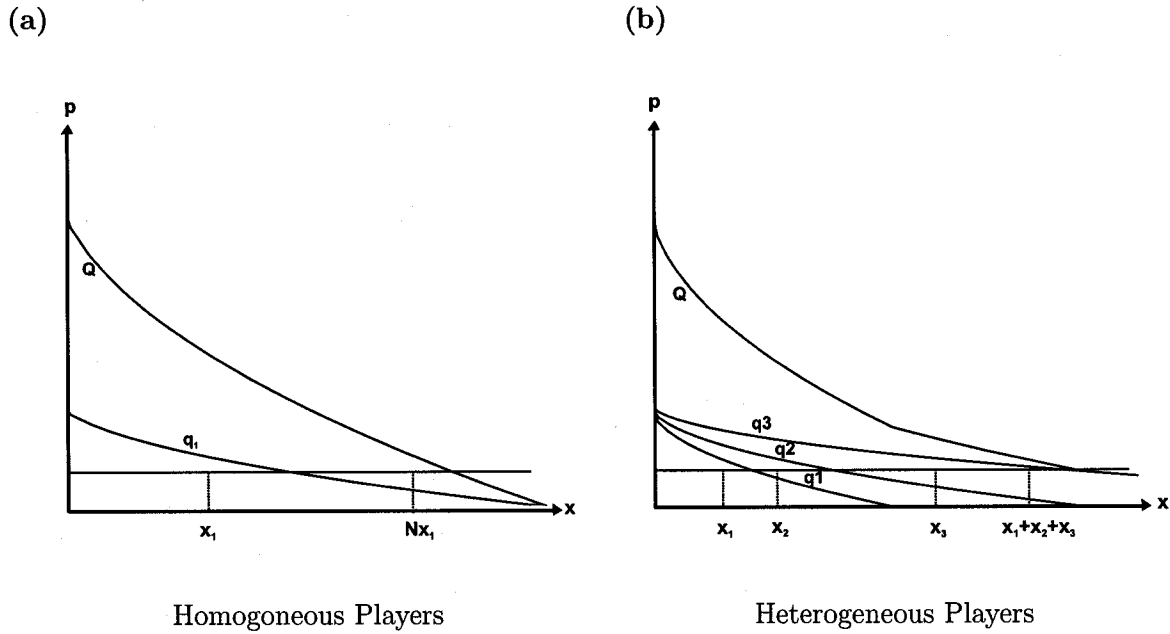


Figure 4.1: Aggregate Demand

The aggregate demand function is defined as  $\bar{Q} = \sum_i^N Q_i(p, s)$  where  $Q_i$  denotes the demand of buyer  $i$  where there are  $N$  buyers. Each  $Q_i(p, s)$  is defined as before:

$$Q_i^*(p, s) = [h(p + s(1 - F(\bar{D}_i)))]_+$$

where  $h(x_i; s, D_i) = \{Q_i \in \mathbb{R}_+ : \bar{u}_i(Q_i; s, D_i) = x_i\}$  and  $\bar{u}_i(Q_i; s, D_i) = \int_{\bar{D}_i(s, Q_i)}^{\infty} u_{Q_i}(Q_i, D_i) dF(D_i)$ .

In Chapter 3, we established the seller-optimal pricing rules as a function of the seller's capacity. In practice, the introduction of multiple buyers increases the probability that the seller's capacity constraint is binding (in so much that it increases the total demand). Whereas in the case of a binding capacity constraint and a single buyer the shortage costs were easily computed, as the difference between the buyer's utility at the unconstrained prices  $u(Q^*(p^*, s^*), s^*)$  and that achieved at the constrained prices and delivery of the seller's total capacity  $u(K, s_C^*)$ , the shortage costs with multiple buyers depend on the rationing rule in place. As discussed, two commonly assumed forms of rationing in the literature are proportional rationing and efficient rationing.

## 4.3 Multiple Sellers

### 4.3.1 Two Sellers with Complete Information

The introduction of a second seller introduces the possibility of strategic interaction between sellers, or players, which we model as a price-setting game with  $k_i \in \{k_L, k_H\}$ , where  $k_i$  denotes seller  $i$ 's capacity and  $0 \leq k_L \leq k_H \leq 1$ . The two sellers compete to serve market demand  $Q$ . We assume first that the sellers know each other's capacity levels, and, hence, we are in a complete-information setting. We impose the following simplification: excess demand either realizes low ( $= 0$ ) or high ( $= 1$ ). Under this simplifying assumption, the number of options that the seller expects the buyer to execute, given that he purchases  $Q$ , is simply  $y = \theta Q$ , where  $\theta = P(D = 1)$ .

As in the example presented in Chapter 3, we assume a specific form of the buyer's utility function. The utility function is the quadratic shortage cost function

$$u(Q, D) = -c(Q - D)^2, \quad (4.1)$$

where  $c$  is some positive weight on the quadratic term,  $Q$  is the number of options held by the seller, and  $D$  is the realization of excess demand  $\tilde{D}$ . The buyer's expected utility is:

$$U(s, Q) - pQ = \theta(-c(Q - 1)^2 - sQ) - pQ. \quad (4.2)$$

The buyer's demand is then:

$$Q(r) = \left[1 - \frac{r}{2\theta c}\right]_+ \quad (4.3)$$

where  $r = p + \theta s$  is the expected price per option from the buyer's perspective. The buyer equates the cost of an option to  $r$ . The quantity served by player  $i$ , denoted as  $q_i$  is defined below:

$$q_i(r_i, r_j; k_i, k_j) = \begin{cases} \min\{1 - \frac{r_i}{2\theta c}, k_i\}, & \text{if } r_i < r_j, \\ \min\{k_i(1 - r_i)/K, k_i\}, & \text{if } r_i = r_j, \\ \min\{\max\{0, 1 - \frac{r_i}{2\theta c} - k_j\}, k_i\}, & \text{if } r_i > r_j. \end{cases}$$

where  $r_i = p_i + \theta s_i$  is the rate that seller  $i$  sets, and  $r_j = p_j + \theta s_j$  is the rate that his opponent  $j$  sets, and  $K = k_1 + k_2$  is the total capacity in the system.

Payoffs to seller  $i$  are:

$$V_i(r_i, r_j; k) = \begin{cases} r_i \min\{1 - \frac{r_i}{2\theta c}, k_i\}, & \text{if } r_i < r_j, \\ r_i \min\{k_i(1 - r_i)/K, k_i\}, & \text{if } r_i = r_j, \\ r_i \min\{1 - \min\{Q(r_j), k_j\} - \frac{r_i}{2\theta c}, k_i\}, & \text{if } r_i > r_j. \end{cases}$$

The following result characterizes the equilibria in capacity-constrained Bertrand competition with complete information.

**Proposition 4.1** *For  $k_H \geq 1$ , the unique pure strategy equilibrium is  $r_i = r_j = c$ , with profits equal to zero. For  $k_L < k_H \leq \frac{1}{4}$ , the unique pure strategy equilibrium is equal to  $r_i = r_j = (1 - k_H - k_L)2\theta c$ , with positive profits. For all other capacity pairs, the equilibrium is in mixed strategies.*

*Proof.* See Appendix. ■

First we show that there is no pure-strategy equilibrium where  $r_i \neq r_j$ . We suppose that there is an equilibrium at  $r_i > r_j$  and show that this cannot be the case. By a symmetric argument, we cannot have  $r_j > r_i$ . Hence we conclude that  $r_i = r_j$ . The two sellers set the same price when they are in the low-capacity and high-capacity states, respectively.

The capacity regions associated with different equilibria in the complete information setting and given below are depicted in Figure 4.2:

R1.  $k_H \geq k_L \geq 1$ : Symmetric equilibrium with  $r_i = r_j = 0$ .

R2.  $k_L < 1$  and  $k_H \geq \frac{1}{4}$ : No pure strategy equilibrium.

R3.  $k_L < k_H \leq \frac{1}{4}$ : Symmetric equilibrium with  $r_i = r_j = (1 - k_H - k_L)2\theta c$ .

The regions in Figure 4.2 correspond to the “over-capacity” (R1) and “under-capacity” (R3) states identified in O&P. The conditions can also be summarized more generally in terms of the demand function  $Q$ , as follows:

C1.  $Q(0) \leq k_L$ : Pure strategy equilibrium at  $(0, 0)$ .

C2.  $Q(0) < L$  and  $q_i(r) = 1 - k_L - \frac{r}{2\theta c} < H$ : Mixed strategy equilibrium.

C3.  $q_i(r) = 1 - k_L - \frac{r}{2\theta c} \geq k_H$ : Pure strategy equilibrium at  $(r(k_H + k_L), r(k_H + k_L))$ .

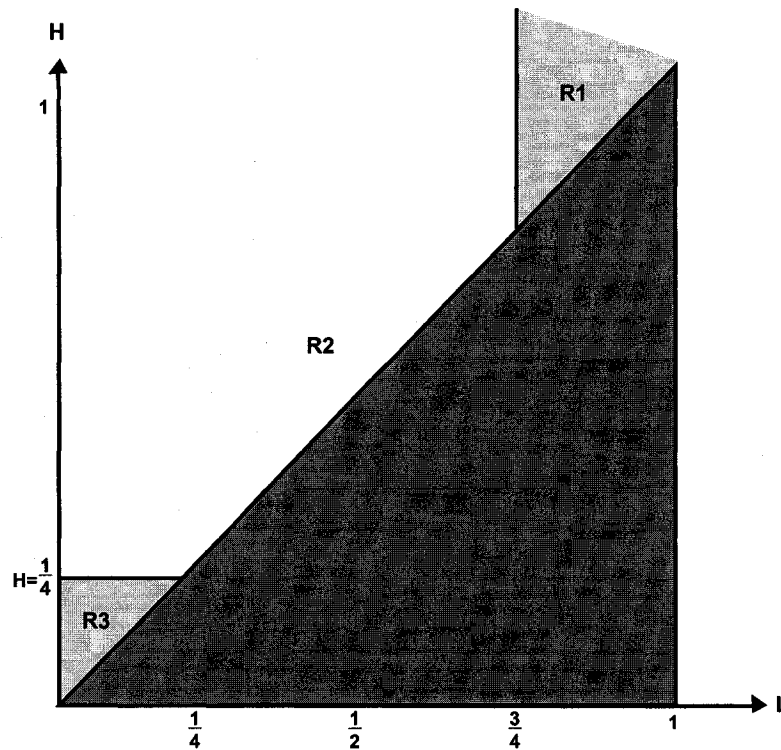


Figure 4.2: Equilibrium Regions: Complete Information

where, in our example above,  $Q(0) = 1$ . As O&P observe, condition C1 renders the capacity constraints irrelevant, and the price-setting competition reduces to the standard Bertrand model, wherein the competitive outcome is achieved (under symmetric costs). That is, both sellers name a price equal to their marginal cost, or reserve value. Under C2, O&P demonstrate how to construct mixed strategy equilibrium (or, in degenerate cases, equilibria). Finally, when there is under-capacity, in C3, the sellers both charge a price at which total capacity clears.

**REMARK 7** If we replace the simplified demand structure above with one where, for instance, demand has a number of possible realizations in the interval  $(0, 1)$ , we will no longer be able to write  $r = p + \theta s$  as the effective price per unit. To see this, consider the case with just one additional level of demand,  $D \in \{0, \frac{1}{2}, 1\}$ . Assume that the probability of being in any one state is equal, with  $\theta = \frac{1}{3}$ . Recall that the buyer's second-period utility of options  $Q$  is  $u(z, D) = -c(z - D)^2$ , where  $z$  is the number of options that he executes, out of a possible  $Q$ . Then the buyer's expected payoff in the first-period is given as:

$$U(s, Q) - pQ = \theta(-c(1-Q)^2 - sQ) + \theta \left( -c \left( \frac{1}{2} - \min \left\{ Q, \frac{1}{2} - \frac{s}{2c} \right\} \right)^2 - s \left( \min \left\{ Q, \frac{1}{2} - \frac{s}{2c} \right\} \right) \right) - pQ$$

Writing  $\min\{Q, \frac{1}{2} - \frac{s}{2c}\}$  as  $\frac{1}{2} - \frac{s}{2c} - (\frac{1}{2} - \frac{s}{2c} - Q)_+$ , we have

$$\frac{\partial}{\partial Q}\{U(s, Q) - pQ\} = \theta(2c(1 - Q) - s) + \theta\left(\delta(d)\left(2c\left(\frac{1}{2} - Q\right) - s\right)\right) - p$$

where  $\delta(d) = 1$  if  $d > 0$  and  $\delta(d) = 0$  if  $d < 0$ , and  $d = Q - \frac{1}{2} - \frac{s}{2c}$ .

Hence

$$Q(p, s) = \left(1 + \frac{1}{2}\delta(d) - \frac{p + \theta s + \theta\delta(d)s}{2\theta c}\right) \frac{1}{1 + \delta(d)}$$

Letting  $r = \theta s + \theta\delta(d)s + p$ , we are still left with the dependence of  $\delta(d)$  on  $s$ .

The seller's payoff is given as

$$V(p, s, Q) = pQ + s\theta Q + s\theta\delta(d)Q + s\theta(1 - \delta(d))\left(\frac{1}{2} - \frac{s}{2c}\right)$$

which we can write as

$$V = rQ + s\theta(1 - \delta(d))\left(\frac{1}{2} - \frac{s}{2c}\right)$$

The dependence of the buyer's purchase *and* exercise decision on the execution decision in the middle demand state, a function of  $s$ , prevent us from simplifying the expressions for  $Q$  and  $V$  in terms of a single price,  $r$ .

Furthermore, with two distinct prices, there is the possibility that the buyer will purchase from both sellers even when capacity constraints are nonbinding. In our analysis above, the buyer *only* purchased from the seller setting the higher price when capacity constraints were binding for the seller setting the lower price. This further complicates the analysis. The model can be numerically implemented to identify states where there is under-capacity (and hence a pure-strategy equilibrium) in, for example, the context of competing irrigation districts selling water to a single buyer.  $\square$

### 4.3.2 Two Sellers with Incomplete Information

Now we assume that the seller's individual capacities are privately held information. In the context of the water market, this scenario captures the possibility that an irrigation district (a seller) may decide not to sell up to its total capacity. The water rights held by a given irrigation district are public information. Hence, total capacity available for sale on the market is known in advance (complete information). However, there remain a number of variables that influence decisions regarding the *actual* supply quantity to be made available on the market. Farmers' planting decisions depend on expected commodity prices,



land rotation practices, and the cost of fallowing acreage, among other things. A district selling water can usually only market up to the amount individual farmers agree to make available by fallowing acreage, where farmers subscribe to a water transfer contract. These district-level decisions are likely to remain private information.

As we show, the prices charged in the incomplete information case are below those charged in the complete information case. More precisely, the price charged by a seller in a high-capacity state (as opposed to a low-capacity state, in a two-capacity state model) is lower than that charged in the incomplete information case. The price charged by the low-capacity seller is equal to the price charged in the complete information case. Hence, buyers are generally better off in the presence of incomplete information. Private information limits the ability of the sellers to coordinate prices, leading the buyer in the high-capacity state to price conservatively.

Let  $k_L, k_H \in [0, 1 - c]$  where  $k_L < k_H$  is fixed. Firm  $i$  has a capacity  $k_i \in \{k_L, k_H\}$ . From the perspective of the other firm, the two capacity realizations are assumed equally likely and independently distributed from the firm's own capacity realization.

An equilibrium strategy in a game of incomplete information identifies an action for each possible state of the world. Specifically, in the two-state pricing game, the equilibrium strategy will be a price pair  $(p_L, p_H)$  corresponding to the respective capacity levels of the firm,  $k_L$  or  $k_H$ . Firm  $i$ 's expected payoff against firm  $j$ 's strategy  $(p_L, p_H)$  is a function of its price  $\pi \in [0, 1]$  and its capacity  $\kappa \in \{k_L, k_H\}$ :

$$\Pi_i(\pi, \kappa) = \frac{\pi - c}{2} (Q_i((\pi, p_L), (\kappa, k_L)) + Q_i((\pi, p_H), (\kappa, k_H)))$$

We restrict attention to the capacity-constrained region  $k_L + k_H \leq 1 - c$  and observe that in this case the equilibrium prices of the two firms must differ. For  $p_H = p_L = 1 - k_L - k_H \geq 1 - c$  there is an incentive for firm  $i$  to "overcut" price when  $k_i = k_H$ . Firm  $i$  receives a strictly greater profit under the overcutting strategy:

$$\Pi_i(1 - 2k_L, k_H) = (1 - 2k_L - c)(k_H + k_H)/2 > (1 - k_L - k_H - c)k_H$$

We first posit that there is an equilibrium in which the high-capacity firm outcharges the low-capacity firm, hence  $p_H > p_L$ . If firm  $j$  plays such a strategy, then firm  $i$  can expect at its equilibrium price  $\pi = p_H$  in state  $k_H$  to tie the high-capacity firm  $j$  and overcut the low-capacity firm. The payoffs are then given as:

$$\Pi_i = \frac{\pi - c}{2} (\min\{(1 - \pi)/2, k_H\} + \min\{\max\{0, 1 - \pi - k_L\}, k_H\})$$

For any given  $k_H \in [k_L, 1 - c - k_L]$ , the last function has the unique maximizer:

$$\pi_H^* = \begin{cases} 1 - k_L - k_H & \text{if } k_H \in [k_L, (1 - k_L - c)/3] \\ (1 + k_H - k_L + c)/2 & \text{if } k_H \in [(1 - k_L - c)/3, (1 + k_L - c)/5] \\ 1 - 2k_H & \text{if } k_H \in [(1 + k_L - c)/5, ((1 - c)/2 + k_L/3)/2] \\ (1 + c)/2 - k_L/3 & \text{if } k_H \in [((1 - c)/2 + k_L/3)/2, 1 - c - k_L] \end{cases}$$

Note that price  $\pi_H^*$  is increasing in capacity.

If  $\kappa = k_L$ , then firm  $i$  expects at its (equilibrium) price  $\pi = p_L$  to tie with the low-capacity firm and overcut the high-capacity firm  $j$ , so that its payoff (as a function of  $\pi$ ) is given as:

$$\Pi_i(\pi, k_L) = \frac{\pi - c}{2} (\min\{(1 - \pi)/2, k_L\} + \min\{\max\{0, 1 - \pi - k_H\}, k_L\})$$

For any  $k_L \in [0, 1 - c - k_H]$ , the last function has an interior maximizer:

$$\pi_L^* = \begin{cases} 1 - 2k_L & \text{if } k_L \in [0, (1 - c)/6] \\ (1 + c)/2 + k_L & \text{if } k_L \in [(1 - c)/6, (1 - c)/4] \\ 1 - k_L & \text{if } k_L \in [(1 - c)/4, 1 - c - k_H] \end{cases}$$

Under this solution, the assumption  $p_H^* > p_L^*$  does not hold, e.g.,  $p_H^* = 1 - k_L - k_H < 1 - 2k_L = \pi_L^*$ . By inspection,  $\pi_H^* \leq \pi_L^*$  for any  $k_L, k_H$  with  $k_L + k_H \leq 1 - c$  and  $k_L \leq k_H$ . We conclude it is therefore not possible that the high-capacity firm outcharges the low-capacity firm.

We now posit the contrary, assuming  $p_L^* > p_H^*$ . The payoff to firm  $i$  when  $\kappa = k_H$ , as a function of the price charged  $\pi$ , is given as:

$$\Pi(\pi, \kappa_H) = \frac{\pi - c}{2} (\min\{(1 - \pi)/2, k_H\} + \min\{1 - \pi, k_H\})$$

where the two firm's tie in the high-capacity state, and firm  $i$  undercuts the low-capacity firm  $j$  otherwise.

For any given  $k_H \in [k_L, 1 - c - k_L]$ , the last function has the unique maximizer:

$$\pi_H^* = \begin{cases} 1 - 2k_H & \text{if } k_H \in [k_L, (1 - c)/6] \\ (1 + c)/2 + k_H & \text{if } k_H \in [(1 - c)/6, (1 - c)/4] \\ 1 - k_H & \text{if } k_H \in [(1 - c)/4, 1 - c - k_L] \end{cases}$$

Note that price  $\pi_H^*$  is increasing in capacity.

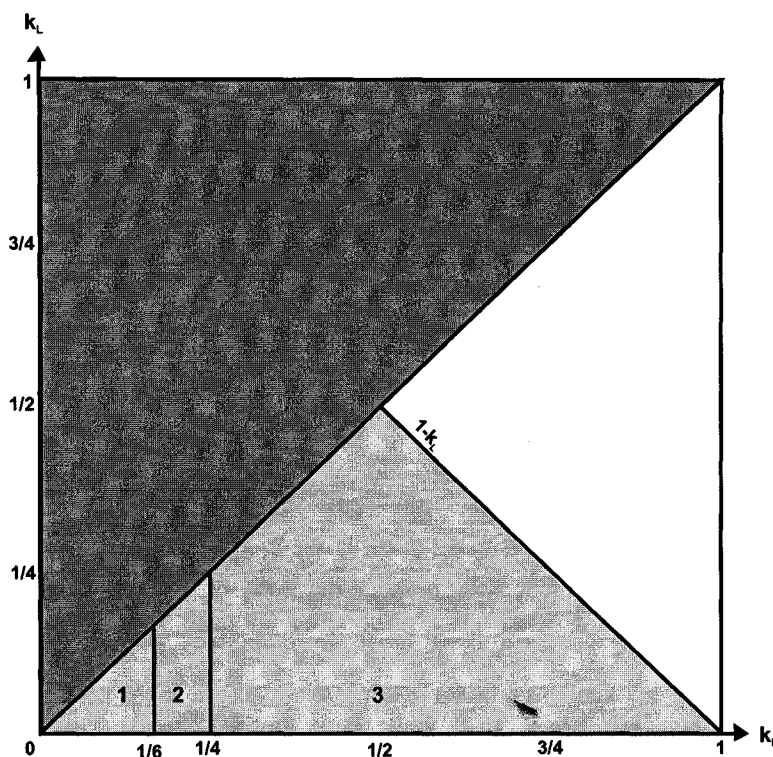


Figure 4.3: Equilibrium Price Regions ( $\pi_H^*$  with  $c = 0$ )

If  $\kappa = k_L$ , then firm  $i$  expects at its (equilibrium) price  $\pi = p_L$  to tie with the low-capacity firm and overcut the high-capacity firm  $j$ , so that its payoff (as a function of  $\pi$ ) is given as:

$$\Pi_i(\pi, k_L) = \frac{\pi - c}{2} (\min\{(1 - \pi)/2, k_L\} + \min\{\max\{0, 1 - \pi - k_H\}, k_L\})$$

For any  $k_L \in [0, 1 - c - k_H]$ , the last function has an interior maximizer:

$$\pi_L^* = \begin{cases} 1 - k_L - k_H & \text{if } k_L \in [0, (1 - k_H - c)/3] \\ (1 + k_L - k_H + c)/2 & \text{if } k_L \in [(1 - k_H - c)/3, (1 + k_H - c)/5] \\ 1 - 2k_L & \text{if } k_L \in [(1 + k_H - c)/5, ((1 - c)/2 + k_H/3)/2] \\ (1 + c)/2 - k_H/3 & \text{if } k_L \in [((1 - c)/2 + k_H/3)/2, 1 - c - k_H] \end{cases}$$

Figures 4.3 and 4.4 depict the pricing regions corresponding to the prices  $\pi_H^*$  and  $\pi_L^*$ , respectively.

If  $\kappa_L < \kappa_H \leq (1 - c)/6$ , then the optimal price charged by the high-capacity player is  $p_H^* = 1 - 2k_H$ . From above, the optimal price charged by the low-capacity player is  $p_L^* = 1 - k_H - k_L$

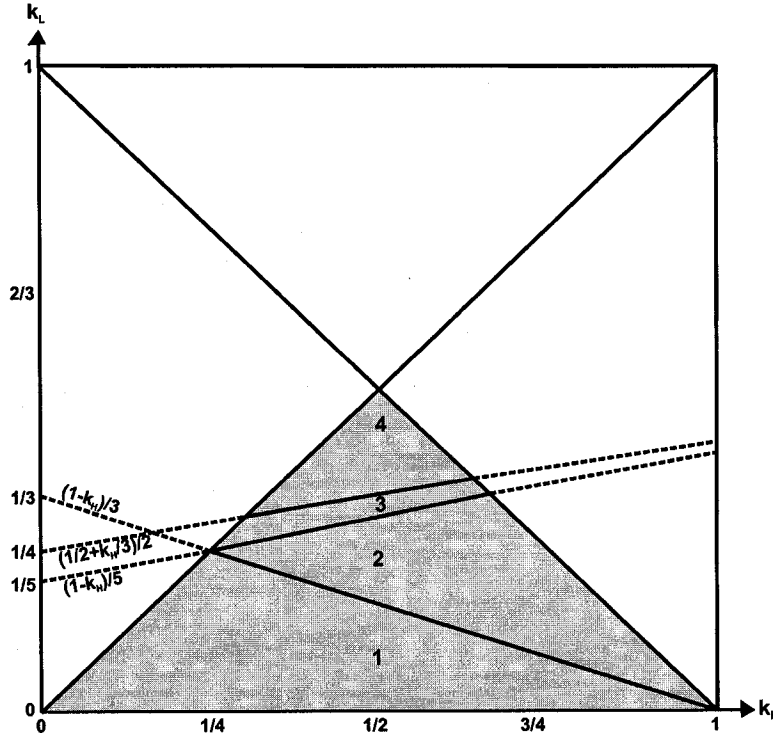


Figure 4.4: Equilibrium Price Regions ( $\pi_L^*$  with  $c = 0$ )

since  $k_L < k_H \leq (1 - c)/6 \Rightarrow k_L \leq (1 - c - k_H)/3$ . Hence our assumption  $p_L^* > p_H^*$  holds for  $k_H \leq (1 - c)/6$ .

If  $2(1 - c)/12 < k_H \leq 3(1 - c)/12$ , then  $p_H^* = (1 + c)/2 + k_H$ . Since we have  $k_L \leq (1 - c)/4 = (1 - c - k_H(1 - c - (1 - c)/4))/3$ , it follows that  $p_L^* = 1 - k_H - k_L$ . It no longer holds that  $p_L^* > p_H^*$ . For example, letting  $c = 0$  and  $2/12 < k_H = 2.5/12 < 3/12$ , we see that  $p_H^* = 8.5/12$  while  $p_L^* = 9.5/12 - k_L < p_H^*$  for  $p_L > 1/12$ . Figure 4.5 illustrates this counterexample, where  $k_H = 2.5/12$ .

**Proposition 4.2** (*Pure-Strategy Equilibrium in an Incomplete Information Setting*) For  $k_L < k_H \leq (1 - c)/6$ , there is a pure-strategy equilibrium with  $p_H = 1 - 2k_H$  and  $p_L = 1 - k_H - k_L$ .

REMARK 8 We can compare the equilibrium prices in a setting with incomplete information to those obtained under complete information. For  $c = 0$ ,  $k_L = \frac{1}{12}$ , and  $k_H = \frac{1}{6}$ , the symmetric pure-strategy Bayes-Nash equilibrium has

$$p_L^* = 1 - k_L - k_H = 1 - (1/12) - (2/12) = 9/12 = 3/4 > 1 - 2k_H = 1 - 2(1/6) = 2/3 = p_H^*$$

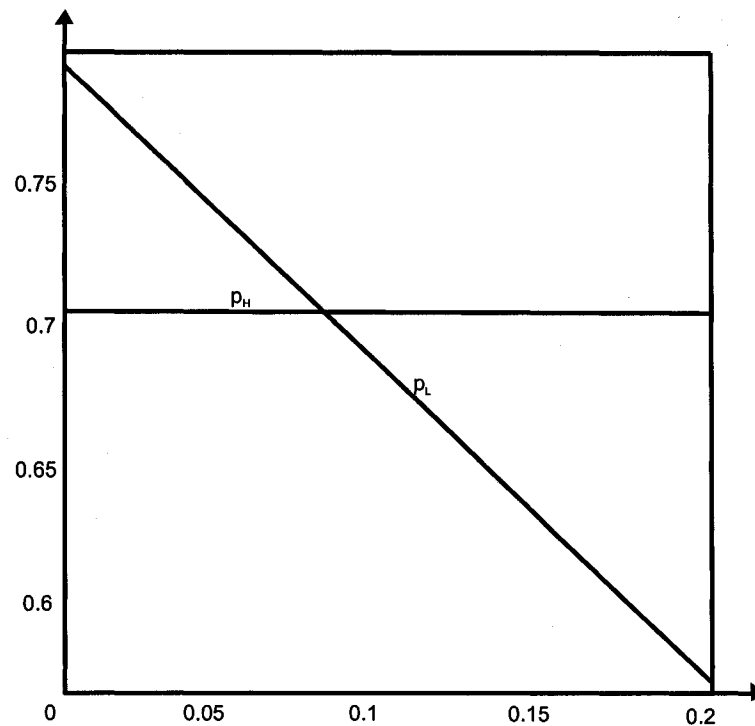


Figure 4.5: Price Comparison:  $p_H^*$  and  $p_L^*$

In the complete information case, the symmetric pure strategy Bayes equilibrium is given by  $p_1^* = p_2^* = 1 - k_1 - k_2 = 3/4$ . It holds that

$$p_1^* = p_2^* \in [p_H^*, p_L^*]$$

□

The buyers are better off in the case with incomplete information, given that  $p_H^* \leq p_1^* \leq p_L^*$ . The price  $p_H^*$  charged in equilibrium by the high-capacity firm is lower than the uniform prices  $p_1^* = p_2^*$  charged by the firms under complete information. At least some fraction of the demand served will be met at this new, lower price.

If capacities are low and the symmetric equilibrium is played in the complete information case, the two firms are essentially extracting monopoly profits. In other words, the presence of two sellers with small capacity relative to demand does not decrease the monopoly power of the sellers. In the high-capacity case when the two sellers play the symmetric equilibrium, the

market power of the firms is indeed reduced. Both firms are charging marginal, implementing the efficient outcome.<sup>2</sup> The buyers appropriate the gains from trade in this scenario.

### 4.3.3 Multiple Sellers with Incomplete Information

The assumption of complete information seems particularly strong as the number of sellers increases. Even if we presume that detailed information about each seller is in the public domain, as is largely the case in the water sector, collecting and analyzing such information for all of one's opponents is likely to be infeasible due to both time and cost constraints. If the players do not know each others' positions perfectly – in particular, do not know each others' capacities – then the problem is one of multilateral contracting with hidden information. Perhaps the most familiar example of this problem is an auction in which there is a single seller, or principal, and multiple buyers, or agents, and asymmetric information regarding valuations.

In the auction context, we revisit the contract being negotiated and assume that the principal and agent(s) sign a forward contract (in place of an option contract). With the arrival of additional sellers, the structure of the marketplace has changed. Seller competition translates into a more favorable position for the buyer, e.g., lower prices. In such a position, the buyer, or principal, may prefer to enter into a forward contract than an option contract. For instance, with a reduced risk of not being able to find a seller (given that there are now many), the buyer may wait and update his demand information before entering into a contract. We suppose that the buyer, having decided on an amount of water for purchase (to meet estimated supply shortfall), conducts an auction to meet this demand.

The auction of interest to us is distinct in two ways. First, the asymmetric information is with respect to (sellers') capacities and not their valuations. As is well known, in multilateral contracting with asymmetric information regarding the sellers' cost and the buyer's value, we are not guaranteed an efficient outcome. The inefficiencies, sometimes referred to as informational rents, arise when the party making the offer inflates his own valuation in hopes of extracting a higher payment from the other party (when he is uncertain of the other party's true valuation). However, we assume here that there is no uncertainty regarding valuations. In the context of water supply, the cost of substitute water for urban supply is generally known. This is due both to the limited supply options available to urban districts in times of shortage and to the fact that such information is disclosed and widely available in

---

<sup>2</sup>This result holds under our assumption of constant returns to scale of the production function in Bertrand competition.

public reports. Facing a short-run supply shortage, an urban water district has very limited alternative supply options. If there is supply available from another water intermediary, it will be at a set rate (often a higher rate referred to as a “penalty rate”). The use of desalination may be an option, if there is a facility online. Again, the cost of which is known. Finally, the option may be to ration, in which case the value to the buyer is the avoided shortage cost of the urban user. While such a shortage cost may not in itself be perfectly known, the intermediary’s assessment of this cost is disclosed in public reports. Finally, we note that the assumption that the buyer’s valuation is known implies that the intermediary also knows the outcome of demand, e.g., knows whether he is in a state of shortage or not.

The asymmetric information in the capacity-dependent auction, while influencing seller competition, does not impact efficiency. With the buyer’s valuation as known, the outcome is efficient, e.g., trade will take place whenever the buyer’s reserve value surpasses the sellers’ reserve values. In context of the discussion in Chapter 3, the auction with known valuation is equivalent to setting an allocation in the second period (as opposed to the first period), once the demand uncertainty has realized.

Under the efficient outcome, we know that total surplus is maximized. We are also interested in the distribution of this surplus. In particular, the principal’s decision regarding a contracting mechanism may hinge on the the portion of the gains from trade that he expects to appropriate under the mechanism in question. From a social welfare standpoint, there may be specific distributional objectives under consideration. The distribution of the surplus in the capacity-dependent auction is a function of the degree of seller competition. Imagine the case with a single seller and known valuations. The situation is that considered in the previous chapter, only the buyer’s value is fixed (and known). Hence the seller simply sets a price to extract the entire surplus, making an offer to charge the buyer his true value (which the buyer will then pay). As competition increases, e.g., as the number of sellers increases, the buyer faces a more favorable distribution of the surplus. We discuss this in more detail after introducing the model of a capacity-dependent auction. In Figures 4.8 and 4.9, we illustrate the shift in the cdf of the buyer’s and seller’s surplus, from  $N = 2$  to  $N = 3$ .

The following multi-unit auction model, a model of capacity-constrained Bertrand price competition, is from Weber (2007): there are  $N$  players and demand is assumed to be fixed and denoted as  $\bar{Q}$  (which can be normalized to 1). The vector of the sellers capacities is given as  $k = (k_1, \dots, k_N)$  and the vector of firms prices is given as  $p = (p_1, \dots, p_N)$ . The payoff

to player  $i$  is given as a function of the price vector  $p$  and the capacity vector  $k$

$$\pi_i(p, k) = \begin{cases} \min \left\{ 1, \frac{1 - \sum_{j \neq i: p_j < p_i} k_j}{k_i + \sum_{j \neq i: p_j = p_i} k_j} \right\} k_i p_i & \text{if } p_i \leq r \\ 0 & \text{otherwise} \end{cases} \quad (4.4)$$

Player  $i$ 's expected profits from a bid below  $r$ , the seller's reserve value, is given as

$$\bar{\pi}(p_i; k_i) = p_i G(\varphi(p_i); k_i), \quad (4.5)$$

where

$$G(s; k_i) = E_{k_{-i}} \left[ \min \left\{ k_i, 1 - \sum_{j \neq i} 1_{k_j > s} k_j \right\} | p_i, k_i \right], \quad (4.6)$$

where  $\varphi(\cdot)$  is the inverse of the equilibrium bidding function, which we denote as  $\beta(\cdot)$  (for  $p_i < r$ ). The FOC for player  $i$  at capacity level  $k_i$  is given as

$$-\frac{\beta'(k_i)}{\beta(k_i)} = \frac{g(k_i; k_i)}{G(k_i; k_i)}, \quad (4.7)$$

where  $g(s; k_i) = \frac{\partial G(s; k_i)}{\partial s}$ . The solution to this ODE is given as follows:

$$\beta(k_i) = r \exp \left[ - \int_{1/N}^{k_i} \frac{g(s; s)}{G(s; s)} ds \right], \quad (4.8)$$

which has an explicit solution for  $N = 2$  (Weber 2008).

The bidding function for  $N = 2$  is illustrated in Figure 4.6. For  $N > 2$ , the optimal bidding function can be approximated numerically. (For  $N = 2$ , the conditional probability distribution of  $\min \left\{ k_i, k_i, 1 - \sum_{j \neq i} 1_{k_j > s} k_j \right\}$  admits a straight-forward representation; for  $N > 2$  this is not the case.) Below we investigate the distribution of the buyer and seller surplus as a function of  $N$  in a numerical implementation of the capacity-dependent auction.

#### 4.3.4 Distribution of Gains from Trade

As mentioned, the buyer is clearly interested in the distribution of the gains from trade associated with a given contracting mechanism. In general, from a social welfare standpoint, we are concerned not only with efficiency but also with distributional objectives. While we know that in the capacity-dependent auction with known valuation the efficient outcome is achieved, the distributional issue is unresolved. The distribution of the gains from trade depends in general on the parameterization of the seller's competition and in particular on the number of sellers and the probability density function of their respective capacities.



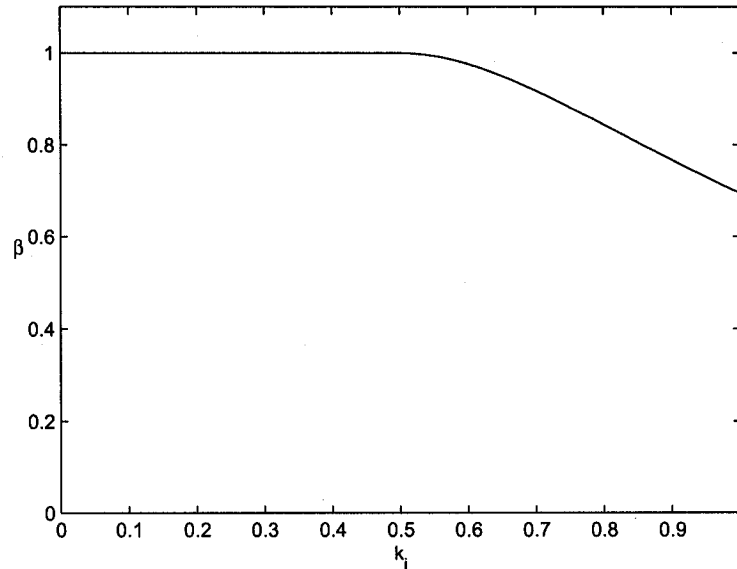


Figure 4.6: Optimal Bidding Function (for  $N = 2$ )

In order to address the question of distribution, we implement the auction numerically, assuming that the sellers' reserve values are zero and the buyer's demand is normalized to  $\bar{Q} = 1$ . For a given  $N$  we can, using simulation techniques, numerically construct the cdf of the buyer's and seller's surplus. The distribution of the buyer's surplus is more favorable in a FOSD sense as  $N$  increases. As there is always some possibility that the sellers' capacities all fall below  $1/N$ , there is also some possibility that the sellers will appropriate all of the surplus. This probability tends to zero for large  $N$ .

We can bound the likelihood of a given surplus loss by one of the two parties using the cdf. This approach is used, for example, in the analysis of computational game theory models, although largely in the context of efficiency loss versus distributional objectives in multi-player games with selfish behavior. (For a detailed discussion see, e.g., the recent book by Nisan, et al. (2007) on algorithmic game theory.) In these models, an interest in efficiency (vs. distributional) losses has prompted the study of the "price of anarchy" and the development of both numerical and analytical models to bound such loss in multi-player games with selfish behavior.

Figure 4.7 below illustrates the optimal bidding function for player  $i$  as a function of the number of total players, from  $N = 2$  to  $N = 7$ . The optimal bid decays rapidly as  $N$  increases. As the buyer's surplus is decreasing in the bid, it is no surprise to observe the

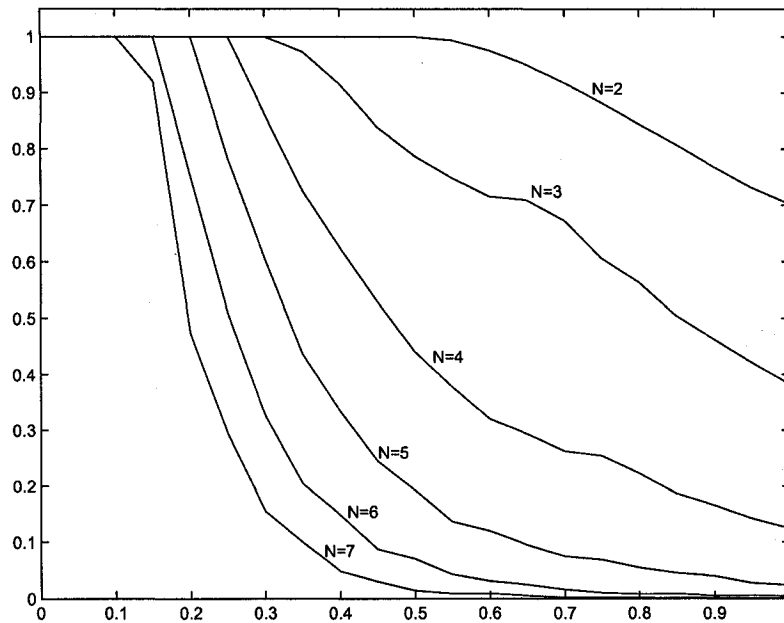


Figure 4.7: Optimal Bidding Function (for  $N \in \{2, \dots, 7\}$ )

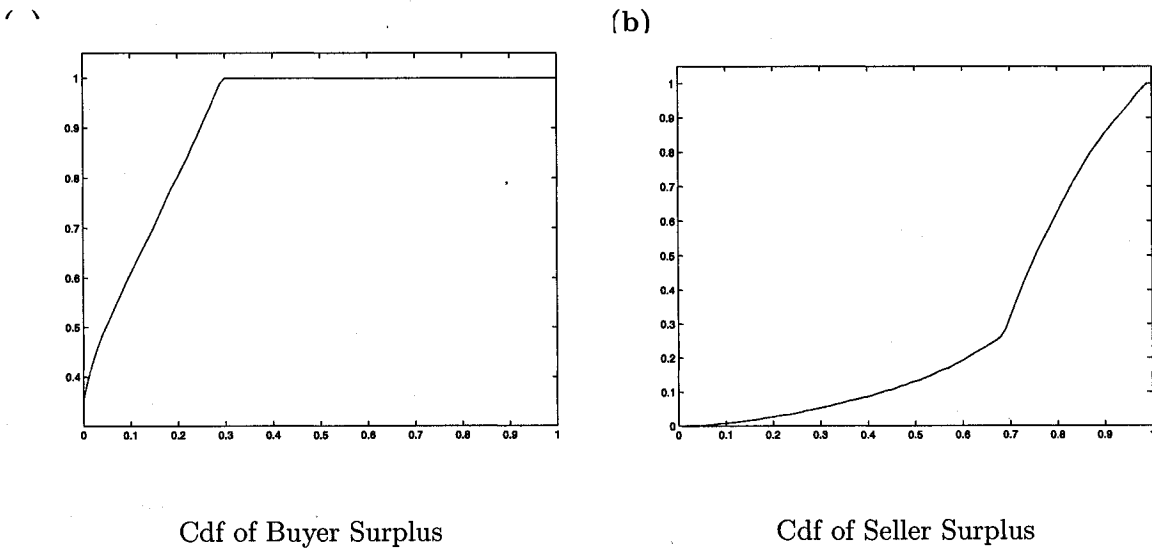
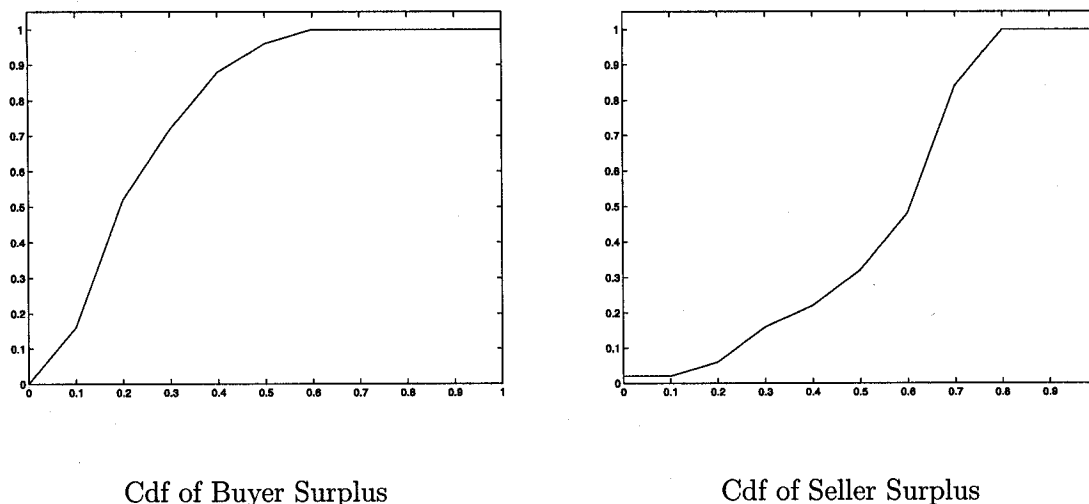


Figure 4.8: Cdf of Surpluses (for  $N = 2$ )

Figure 4.9: Cdf of Surpluses (for  $N = 3$ )

FOSD shift in the distribution of the buyer's surplus, from Figure 4.8 to Figure 4.8.

Figure 4.8 illustrate the cdfs of the buyer's and seller's surpluses, respectively, in the two-player case. Figure 4.9 illustrate the adjustment in cdfs of the buyer's and seller's surpluses, respectively, in the three-player case. With two sellers, the buyer's surplus is below 30% with probability one. With the addition of a third seller, the surplus is now below 60% with probability one (there is a 75% chance that the surplus is below 30% and, hence, a 25% chance that is above this and above the highest possible surplus achievable in the two-player game). The sellers' market power erodes rapidly with the entrance of additional sellers.

## 4.4 Summary

The presence of the additional sellers in the contracting market significantly changes the option contracting model, necessitating the treatment of strategic interactions between sellers. The price-setting policy of the individual sellers adjusts and hence, the final contract struck between the buyer and the seller (or sellers, as may be the case when the seller offering the lowest price does not have sufficient capacity to meet total demand) is revised. In this chapter, we have emphasized the influence of capacity constraints on the outcome with multiple sellers. Without capacity constraints and with symmetric costs, Bertrand price competition between two or more sellers results in the competitive outcome. With capacity constraints,

we see that sellers are able to appropriate some of the gains from trade. The efficiency of the outcome depends on the informational asymmetries regarding valuations (not capacities).

The impact of capacity constraints on the model with multiple sellers can be interpreted in terms of the sellers' market power. In the case of a single seller, as seen in Chapter 2, pricing is monopolistic and the seller appropriates a larger share of the total surplus than he would do under competitive pricing. There are associated efficiency losses (deadweight losses) under monopolistic pricing. The introduction of additional sellers represents the erosion of this market power. However, capacity constraints reinstate market power by creating a form of independence between sellers. This independence, where if the buyer purchases from one seller he will still likely purchase from another, can be interpreted as a reinstatement of market power. It is therefore not surprising to find that the seller appropriates a portion of the total surplus in the presence of capacity constraints – his surplus is an increasing function of his market power. This market power, in turn, is a decreasing function of the number of sellers  $N$ , and, as discussed in the previous section, the distribution of the buyer's and sellers' surplus depends critically on  $N$ .

In both extensions of the contracting model – Bertrand price competition, where two sellers announce prices simultaneously, and a single-principal multi-agent capacity-dependent (or multi-unit) auction – we have assumed that valuations are known. Under these assumptions in the capacity-dependent auction, we achieve the efficient allocation. Whether the buyer would be willing to institute a capacity-dependent auction, given its potential for efficient allocation, may depend on the anticipated distributional outcome. The possible distributional outcomes can be further investigated in a numerical model, as demonstrated in the previous section. As the number of sellers increases, the buyer expects to appropriate a larger share of the gains from trade. The fact that there may not be sufficient underlying capacity to satisfy total demand is not of course remedied by conducting an auction. Concern regarding this fact might lead the buyer to consider incentives for seller participation in the auction.

## 4.5 Appendix

### Proof of Proposition 4.1.

Suppose that  $r_i > r_j$  is an equilibrium.

If  $r_i > r_j$  is an equilibrium, then,  $r_i = \arg \max_{r_i} \{r_i \min\{1 - \min\{Q(r_j), k_j\} - \frac{r_i}{2\theta c}, k_i\}\}$  must be a best response for player  $i$ , where

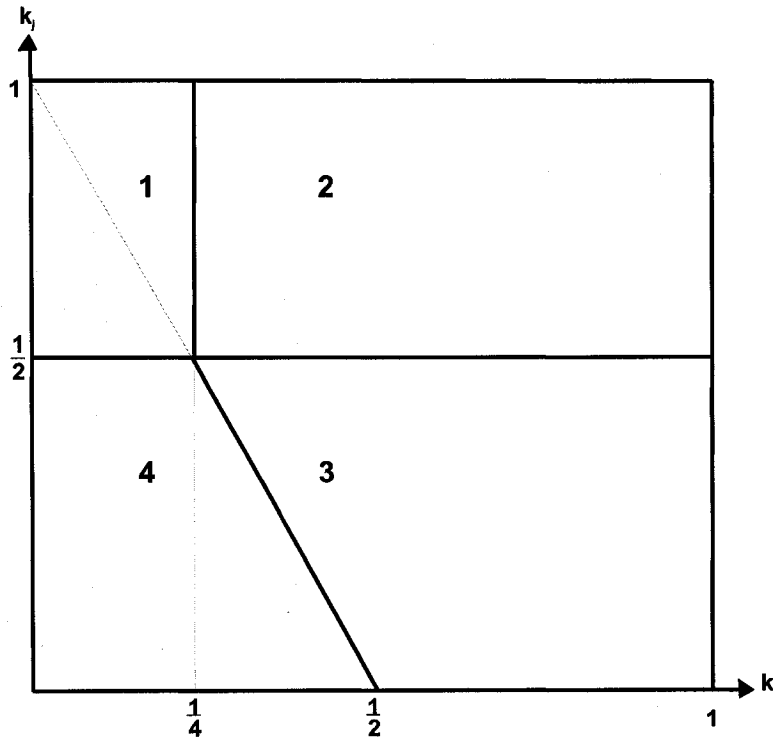
$$r_i = \begin{cases} \frac{1}{2}r_j & \text{if } k_j \in [1 - \frac{r_j}{2\theta c}, 1) \text{ and } k_i \in [\frac{1}{2}\frac{r_j}{2\theta c}, 1) \\ r_j - 2\theta ck_i & \text{if } k_j \in [1 - \frac{r_j}{2\theta c}, 1) \text{ and } k_i \in (0, \frac{1}{2}\frac{r_j}{2\theta c}) \\ (1 - k_j)\theta c & \text{if } k_j \in (0, 1 - \frac{r_j}{2\theta c}) \text{ and } k_i \in [\frac{1-k_j}{2}, 1) \\ (1 - k_j - k_i)2\theta c & \text{if } k_j \in (0, 1 - \frac{r_j}{2\theta c}) \text{ and } k_i \in (0, \frac{1-k_j}{2}) \end{cases}$$

and, likewise,  $r_j = \arg \max_{r_j} \{r_j \min\{1 - \frac{r_j}{2\theta c}, k_j\}\}$  must be a best response for player  $j$ , where

$$r_j = \begin{cases} \theta c & \text{if } k_j \in (0, \frac{1}{2}] \\ (1 - k_j)2\theta c & \text{if } k_j \in (\frac{1}{2}, 1) \end{cases}$$

If  $k_j \in [\frac{1}{2}, 1)$ , then  $r_j = \theta c$ , and either  $k_i \in [\frac{1}{4}, 1)$ , since  $\frac{1}{2}\frac{r_j}{2\theta c} = \frac{1}{4}$ , or  $k_i \in (0, \frac{1}{4})$ . These are the first two cases. Otherwise, if  $k_j \in (0, \frac{1}{2})$ , in which case, either  $k_i \in [\frac{1-k_j}{2}, 1)$  or  $k_i \in (0, \frac{1-k_j}{2})$ . These are the four regions depicted in Figure 4.10.

- Case 1. If  $k_j \in [\frac{1}{2}, 1)$  and  $k_i \in [\frac{1}{4}, 1)$ , then  $r_j^* = \theta c$  and  $r_i^* = \frac{1}{2}\theta c$ . But this is a contradiction, as we have assumed that  $r_i > r_j$ . Hence this cannot be an equilibrium.
- Case 2. If  $k_j \in [\frac{1}{2}, 1)$  and  $k_i \in (0, \frac{1}{4})$ , then  $r_j^* = \theta c$  and  $r_i^* = r_j - 2\theta ck_i = (1 - 2k_i)\theta c$ . Again, this is a contradiction:  $r_i < r_j$  for  $k_i < \frac{1}{4}$ , and we have assumed that  $r_i > r_j$ . Hence this cannot be an equilibrium.
- Case 3. If  $k_j \in (0, \frac{1}{2})$  and  $k_i \in [\frac{1-k_j}{2}, 1)$ , then  $r_j^* = (1 - k_j)2\theta c$  and  $r_i^* = (1 - k_j)2\theta c$ . Again, a contradiction:  $r_i^* < r_j^*$ .
- Case 4. If  $k_j \in (0, \frac{1}{2})$  and  $k_i \in (0, \frac{1-k_j}{2})$ , then  $r_j^* = (1 - k_j)2\theta c$  and  $r_i^* = (1 - k_j - k_i)2\theta c$ . Then  $r_i^* < r_j^*$  and again we have reached a contradiction.

Figure 4.10: Capacity Regions for  $r_i > r_j$ 

Since cases one through four are mutually exclusive and exhaustive, we conclude that there is no equilibrium where  $r_i > r_j$ . By a symmetric argument, we cannot have  $r_j > r_i$ .

We examine the remaining possibility:  $r_i = r_j$ . Below we examine the possible cases. In equilibrium, we must have that  $r_i = \operatorname{argmax} \{r_i \min\{\frac{1}{2}(1 - \frac{r_i}{2\theta c}) + (\frac{1}{2}(1 - \frac{r_i}{2\theta c}) - k_j)_+, k_i\}\}$  constitutes a best response for both players  $i$  and  $j$ , where

$$r_i = \begin{cases} \theta c & \text{if } k_i \in [\frac{1}{4}, 1) \\ (1 - k_i)\theta c & \text{if } k_i \in (0, \frac{1}{4}) \text{ and } k_j \in [\frac{1}{4}, 1) \\ (1 - k_i - k_j)2\theta c & \text{if } k_i \in (0, \frac{1}{4}) \text{ and } k_j \in (0, \frac{1}{4}) \end{cases}$$

First of all we observe that if  $k_i \in (0, \frac{1}{4})$  and  $k_j \in (\frac{1}{4}, 1)$  then, by the definition above, sellers  $i$  and  $j$  will actually want to charge *different prices* – a contradiction to the assumption that  $r_i = r_j$ . Therefore, we can eliminate the middle case as a possible equilibrium. We treat the other two possibilities separately: (1)  $k_i > \frac{1}{4}$  and  $k_j > \frac{1}{4}$  and (2)  $k_i \leq \frac{1}{4}$  and  $k_j \leq \frac{1}{4}$ .

Suppose that  $r_i = r_j$  is an equilibrium.

*Case 1.* If  $k_j \in [\frac{1}{4}, 1)$  and  $k_i \in [\frac{1}{4}, 1)$ , then  $r_i = r_j = r = \theta c$ . At this price, both players serve:

$$\frac{1}{2}Q(r) = \frac{1}{2} \left(1 - \frac{r}{2\theta c}\right) = \frac{1}{4}$$

However, as  $k_i \geq k_j > \frac{1}{4}$ , by the undercutting argument both players have an incentive to shave this price. That is we can find an  $\epsilon > 0$  such that:

$$(r - \epsilon)Q(r - \epsilon) > r \frac{1}{2}Q(r)$$

For the sake of concreteness,

$$(r - \epsilon) \left(1 - \frac{r - \epsilon}{2\theta c}\right) = (\theta c - \epsilon) \left(1 - \frac{\theta c - \epsilon}{2\theta c}\right) > r \frac{1}{2} \left(1 - \frac{r}{2\theta c}\right) = \frac{1}{4}\theta c$$

for all  $0 < \epsilon < \sqrt{\frac{1}{2}}\theta c$ .

By a similar argument, at any price  $r > 0$ , there will be an incentive to shave price. For  $r = 0$ , there is no incentive to shave price, since doing so implies a subsidy and a strictly negative payoff for the seller. Below we show that there can be an equilibrium at  $r_i = r_j = r = 0$ , if the capacity levels of the two player are high enough.

As noted, neither player will charge a negative price. We ask then, *is there an incentive for either player to raise price?*

The answer to this question depends on the capacities of the players. If  $L < 1$ , the answer is yes, and, hence, there is no pure strategy equilibrium. To see this, we examine the payoff to player  $i$  if he raises price by  $\epsilon > 0$ :

$$V_i = (r_i + \epsilon) \left(1 - k_j - \frac{r_i + \epsilon}{2\theta c}\right) = \epsilon \left(1 - k_j - \frac{\epsilon}{2\theta c}\right)$$

at  $r_i = 0$ . This payoff is greater than zero whenever

$$1 - k_j - \frac{\epsilon}{2\theta c} > 0$$

For any  $k_j < 1$ , we can find an  $\epsilon > 0$  for which this expression is positive. Hence the condition for there to be a pure strategy equilibrium at  $r_i = r_j = 0$  is  $k_j \geq 1$  and, by symmetry, we must also have  $k_i \geq 1$ . Since  $H > L$ , the condition is simply  $L \geq 1$ .

If this condition is not met, then there can be no pure strategy equilibrium, since players will have an incentive to undercut each other at any prices above zero, and at zero, at least

one player will have an incentive to raise prices.

*Case 2.* If  $k_i \geq k_j$  and  $k_i \in [\frac{1}{4}, 1)$ , then the optimal price is  $r = (1 - k_i - k_j)2\theta c = (1 - H - L)2\theta c$ .

At this price, the total quantity demanded is given as

$$Q(r) = 1 - \frac{r}{2\theta c} = H + L$$

Therefore, player  $j$  serves

$$q_j(r) = \min \left\{ \frac{1}{2} \left( 1 - \frac{r}{2\theta c} \right) - \left( 1 - \frac{r}{2\theta c} - k_H \right)_+, k_L \right\} = k_L$$

and player  $i$  serves

$$q_i(r) = \min \left\{ \frac{1}{2} \left( 1 - \frac{r}{2\theta c} \right) - \left( 1 - \frac{r}{2\theta c} - k_L \right)_+, k_H \right\} = k_H$$

since we have assumed  $k_i \geq k_j$ .

Neither player has an incentive to shave price: at a lower price, each would offload the same quantity but make a strictly lower profit. Increasing price would lead to a reduction in sales, a contradiction to the bindingness of the constraints. We conclude that the price  $r = (1 - H - L)2\theta c$  is an equilibrium.  $\square$





# Chapter 5

## Conclusion

### 5.1 Contributions

This thesis contributes three economic models, motivated by questions regarding the allocation of water and the structure of optimal contracts in emerging water markets, in the face of costly water shortages and uncertain supply and demand. To the best of our knowledge, these questions have not been satisfactorily addressed in the literature to date. They are arguably questions of great importance in light of the rising costs and increasing uncertainty associated with water supply. Below we revisit the three research questions postulated in Chapter 1 and review our conclusions.

**Research Question 1** *Given fixed short-run supply contracts and a given storage level, what is the optimal joint rate-setting and supply procurement policy by a water intermediary facing both supply and demand uncertainty?*

The optimal joint price-inventory control policy under the backlogging assumption is a base-stock-list-price policy. That is, the intermediary optimally places orders whenever the inventory level is below the base-stock level, and there is an optimal price (the “list price”) associated with the base-stock level. When the inventory level exceeds the base-stock level, the intermediary offers discounts. As we show, when the backlogging assumption is relaxed, we are no longer guaranteed the optimality of this policy, as backlogging is needed to guarantee concavity of the payoff function. We can however still find optimal policies given the well-structured nature of our optimization problem.

In the application of the model to the Metropolitan Water District of Southern California (MWD), we investigate the current pricing policy used by MWD. We find that the current rate falls somewhere between the optimal welfare-maximizing and profit-maximizing rate. In general, the optimal order level and price will be higher and lower, respectively, under welfare maximization versus profit maximization. The model generates a number of policy insights, including the direction of price changes (increasing) under increases in the variance of the supply and demand uncertainty and the direction of inventory-price policies in the face of increasing costs.

**Research Question 2** *What is the structure of the optimal bilateral option contract given downstream demand uncertainty, uncertain valuations on outside options for the buyer and the seller, and the possibility of infrastructure failure?*

In Chapter 3, we derived optimal contracting policies under three contracting scenarios. Under the first scenario, the buyer has no outside option. Under the second scenario, the buyer has an outside option available at a fixed price. The presence of an outside option restricts the seller's pricing policy, and we find that for a low enough value of the outside option the seller's optimal policy is to charge a zero upfront fee and to extract the maximum amount from the seller in the second period, e.g., by setting the strike price equal to the value of the outside supply option. In the final scenario, with joint spot market access, we find that in general the upfront fee will be higher and that the seller holds less options.

The presences of demand uncertainty in the contracting model leads to results that differ from those in Wu et al. (2002) for a contracting model with joint spot market access. With demand uncertainty, we can find parameterizations of the buyer's utility function, e.g., shortage cost function, such that contracting is feasible with joint access and no probability of failed delivery on the seller's part. Wu et al. found that without demand uncertainty, the feasible conditions for contracting with joint spot market access required that the seller face a positive probability of not being able to offload his entire supply onto the market.

The study of bilateral option contracts is motivated by the observation that permanent transfers in the water market have been stalled to date, owing both to institutional resistance and high transaction cost, discussed in more detail below. We conclude Chapter 2 with an application of the model to the California water market and an established set of contracts between MWD and one of the largest irrigation districts in the state, Glenn Colusa Irrigation District (GCID). Our results suggest that the prices set in these contracts were too low. Part

of the explanation for this may lie with the fact that the contract was set by negotiation and not the direct result of seller-set prices. Nonetheless, low prices are a concern on two fronts, both because they can lead to inefficiencies in resource use and also because they can sour future transactions. As we discuss below, the options model accurately captures the uncertainties determining option value, and therefore provides insight into sound pricing policy in the future.

**Research Question 3** *How does the entrance of additional buyers and sellers into the contracting market impact the optimal structure of option contracts?*

In Chapter 4, we considered a variation to the option contracting model in which there is a single seller and multiple buyers, and vice versa. We show that an increase in the number of buyers does not change the structure of the contracting model. It does however change the actual prices in so much as aggregate demand will be stronger than individual demand, allowing the seller to set higher prices. We assume that the seller can then serve demand on a willingness-to-pay basis.

The presence of additional sellers in the contracting market significantly changes the structure of the contracting model. The strategic interaction between sellers must be taken into account. We examine the case of duopoly under complete information and under incomplete information in the presence of capacity-constraints. Under a simplifying assumption on the structure of the random demand (demand either realizes high or low, i.e., 0 or 1), the two-price structure reduces to a single tariff. In this case, the analysis by Osborne and Pitchik (1986) applies, and there is a unique pure-strategy equilibrium in both the under-capacity and over-capacity states of the supply system. (In the mid-capacity case, there are mixed strategy equilibria.) The two sellers charge the same price in equilibrium in the under- and over-capacity states. In contrast, in the model with incomplete information and two capacity states (high or low), the seller in the high-capacity state charges a lower price than the seller in the low-capacity state. Furthermore, the price charged by the high-capacity player in the incomplete information state is lower than that charged by the high-capacity player in the complete information state. The presence of private information appears advantageous to the buyer, resulting in lower prices. We conclude Chapter 4 with a discussion of a multi-unit auction. Such auctions may experience increasing interest if the number of willing sellers (irrigation districts) increases and if large buyers such as MWD are in a position to arrange these auctions.

## 5.2 Application of the Models

### 5.2.1 Price-Inventory Control at MWD

In Chapter 2, we discuss the application of the joint inventory-price control model to MWD, California's largest water intermediary. The implementation of such a pricing model would represent a reform in the rate-setting practices of the agency. The current rate-setting process is done under the "cost of service" methodology, in accordance with the American Water Works Association (AWWA) guidelines (AWWA 2000). Rates are set annually and approved by a board of directors up to six months prior to the date on which they come into affect. See Figure 1 below. In Figure 2 we illustrate the change in rate-setting policy under a price-inventory control policy, where rates can adjust based on updated information regarding supply realizations and current storage levels.

The cost of service methodology is a four-step process, as depicted in Figure 3. In the first step, the agency identifies all of its separate project expenses or line-items in the annual budget. This total is the agency's "revenue requirement." The costs are then grouped into cost categories, each associated with a different component of the water supply system. (See Figure 4.) Finally, these cost categories are used a a basis for allocating costs via service rates.

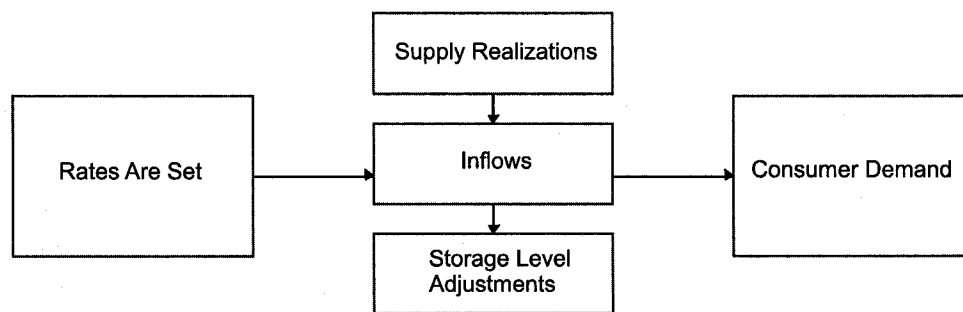


Figure 5.1: Current Rate-Setting Practice

For each cost "block," the agency (1) estimates the total costs associated with activity and (2) attempts to allocate the fees for the block to the responsible user(s). For instance, if one cost block is "supply expansion" and a certain set of users will benefit from the expansion, then that cost block will be charged to those users on a proportional basis, where the proportional basis is intended to reflect the proportional value of the service to the users. The method is intended to approximate marginal cost pricing, with marginal costs generally

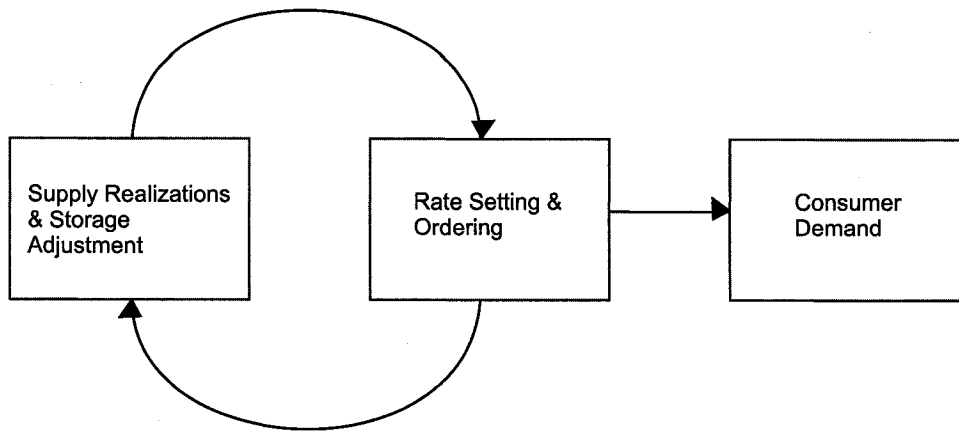


Figure 5.2: Dynamic Rate-Setting Process

estimated as a constant for each block (as the total cost divided by the number of units of water associated with the block).

There is a recognized division between costs incurred for serving “base demand” and costs incurred for “peak demand,” and the agency attempts to allocate these costs separately. The Tier 1 rate is that charged to consumers who incur base demand costs, while the Tier 2 rate is charged to consumers who place higher demands on the system (relative to historic demand). While separate rates are also established for different services (such as treatment and conveyance), since they apply to every unit of water, they essentially comprise a (single) fixed charge per unit of water ordered, or delivered. There are a number of costs allocated as a fixed fee to the member agencies of MWD.

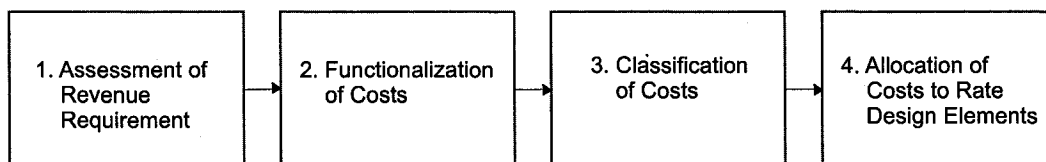


Figure 5.3: Steps in the “Cost of Service” Rate-Setting Methodology

The current rate-setting methodology provokes three critiques. First, the process is fundamentally non-dynamic. It does not take into consideration the fluctuation in storage levels over time. Second, it fails to capture the true impact of demand and supply uncertainty on the system (and the underlying value function of interest). Demand and supply “scenarios”

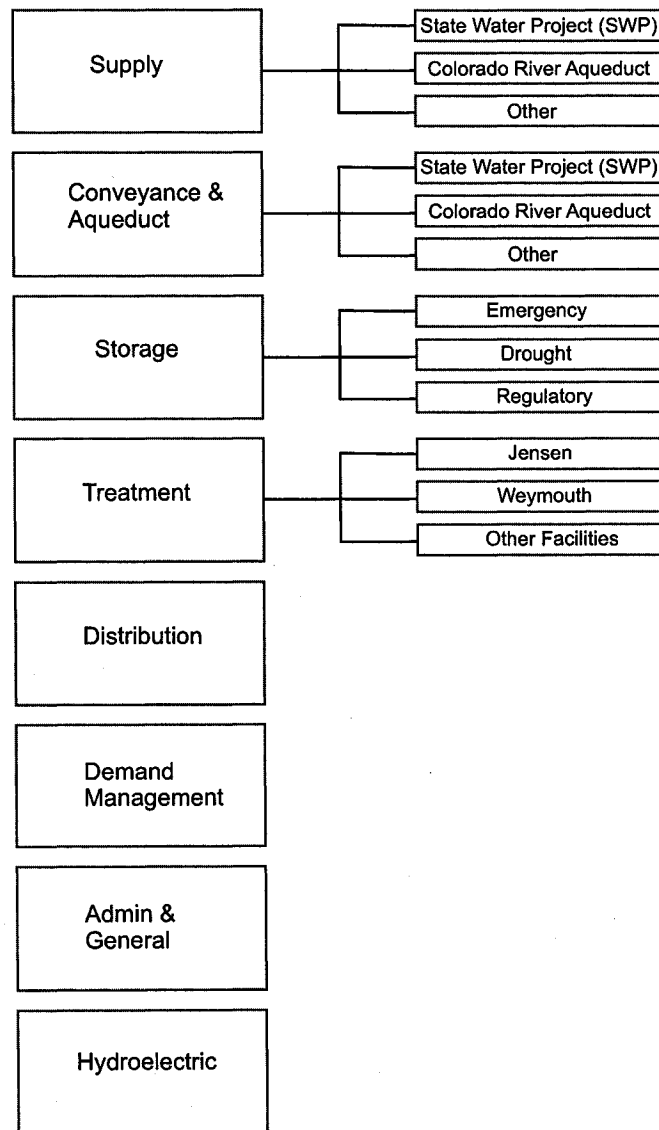


Figure 5.4: Categories for Cost Allocation

are used in the rate-setting process. In some cases, a single scenario of average demand and supply provides the baseline for planning. Use of the average implies an underestimate of the shortage costs (when shortage costs are convex) or an overestimate (when the shortage costs are concave). Third, the rate-setting process does not explicitly recognize the intermediary as having an “ordering decision” that can be adjusted, as would be the case in the presence of flexible supply contracts (such as option contracts).

These three shortcomings lead to non-optimal decision-making and implied efficiency losses. The price-inventory model from Chapter 2 captures the true impact on social welfare of the demand and supply uncertainties, represented as probability distributions *not averages*. This type of pricing model has been actively adopted in the retail industry and can offer similar efficiency gains in the water management sector. The detailed cost estimation models currently applied under the AWWA rate-setting process could be used to provide the necessary cost parameters for input to the joint inventory-price control model.

### 5.2.2 Option Contracting in the California Water Market

As discussed in Chapter 1, the development of a water market in California and elsewhere would facilitate the reallocation of water. Under reallocation, gains from trade arising from shifts from lower-value to higher-value end uses (and the avoidance of costly shortfalls in supply in an urban setting) are realized. The development of such a market depends on a number of institutional, legal, and economic factors.

In California, the existence of literally hundreds of separate water districts with heterogeneous water rights (allocations) and widely disparate prices for water within districts suggests that there are significant gains from trade. Yet a water market in which large volumes of water are freely transferred between districts has not evolved to date. The introduction of temporary transfers and a shift from the historical focus on permanent water rights transfers, or statutory transfers, could stimulate the market.

A review of the institutional and legal setting reveals a number of impediments to the development of the water market, many of which have their origin in the state’s historical water rights system. This system was designed to prevent, rather than facilitate, transfers. The institutional structure set in place by the California legislature in the early 1900s accomplished many of its aims, primary among them to band together local water users to encourage water use for the common good. The legislature formed irrigation districts and water mutuals that, as described in Introduction, work much like municipalities. These districts are overseen by boards, which are in turn elected by the districts’ residents, with



voting rights either held in proportion to land ownership or on a one-member-one vote basis. The districts have been largely successful in promoting equitable and cooperative water use (and avoiding the “tragedy of the commons” associated with public goods). However, their formation also set the stage for future conflicts by (1) creating incentives for district members to block proposed water transfers by other members and (2) failing to establish rules for the allocation of profits from such transfers to members of the district. Water districts generally lack the authority under state statute to disburse proceeds from water sales. Even with this authority, there is not an established protocol for disbursing proceeds among members with different land holdings and different water usage patterns (and yet possibly equal voting, or membership stakes, in the district).

Among the incentives for water district members to block transfers include (1) retention of stature and power, as derived from the control of the district’s valuable water resources (which would otherwise be transferred away), by the district managers, (2) prevention of rate increases, a motivation for district members who might need to buy additional water and don’t wish to pay the (higher) market rate and, hence, hope to keep the rate (artificially) low by blocking transfers and, (3) avoidance of economic losses, such as would be suffered by non-water rights holders, i.e., laborers, fertilizer salesman, millers and crop marketers in an agricultural district, all of whose livelihoods are threatened by water transfers (Thompson 1993).

High transaction costs associated with permanent water rights transfers are another major market friction. Part of the transaction cost can again be related back to the institutional structure. Lacking a well-defined social norm for the distribution of proceeds amongst heterogeneous members of an irrigation district, the negotiation of a deal that is perceived as fair (and hence receives general buy-in) can be a costly process both in terms of time and legal fees. As the number of temporary (and permanent) transfers increases over time, the establishment of an acceptable template, of sorts, for transfers may help lower these costs.

Another contributor to high transaction costs is the necessary environmental impact assessment (EIA) for a permanent transfer. The contracting parties must conduct extensive studies and provide a scientific review of the potential long-term environmental impacts of a permanent water transfer. In addition to the EIA, the law requires that the parties involved in the transfer must perform due diligence, a costly and time consuming process which involves proving that there is no harm to other parties or the environment. (The burden of proof is on the transacting parties to show “no harm” and not on the downstream parties who may claim injury.) More than one attempt to transfer permanent water rights has fallen through after years of investment in the process. Finally, the valuation of the water right itself can

pose an impediment to trade. A water right is a perpetuity, with uncertainty surrounding the quality and the actual volume of the flows (which is dependent on the seniority of the right), not to mention the actual future value-in-use.

Both the institutional resistance and the high transaction costs associated with water transfers are considerably alleviated by a shift from permanent to temporary transfers. As such, option contracts offer a politically and economically viable solution in markets where a number of impediments to the transfer of permanent water rights currently exist, as is the case in California. Resistance on the part of the farming community to water transfers is reduced when the threat to livelihoods is removed, and temporary transfers allow periodic cash infusions which can actually be good for the community. The option contracting models developed in Chapters 3 and 4 provide a basis for pricing these contracts and assessing the efficiency gains obtained under such contracts.

There are a number of additional policy actions that could facilitate the development of such a market, including investment in additional infrastructure, required for the market reallocation between northern and southern users to take place on a scale that can adequately meet predicted future shortages.<sup>1</sup> We touch on three that arose during our work in the California water market.

The first is a technical issue, requiring legislative review, with regard to an equivocal item in the state's water code: if a farmer transfers his surface water and at the same time forgoes use of his groundwater, since he is required to fallow his land and therefore has no need to irrigate it, has he in effect forfeited his groundwater right? As discussed above, standard practice in irrigation districts with aquifers is to pump groundwater, usually from private wells, to augment surface water in dry years. This sets a precedent for groundwater use in dry years. If the groundwater is not used when the land has been fallowed, it is unclear under current statute whether or not a "use it or lose it" clause implies the groundwater right has been abandoned. Clearly this issue is of concern to the farmer, given that a one-term surface water transfer could then result in a loss of groundwater pumping rights and render the land non-arable during future dry periods. A statutory change to the water code is required to resolve this issue.

A second issue for legislative review is that of subsidies. The current structuring of subsidies under the Farm Bill is such that farmers are paid a subsidy per cwt for a given crop. In years that land is not cultivated, the subsidy is foregone. Quite apart from other distortionary

---

<sup>1</sup>This issue is at the center of discussion surrounding the peripheral canal, the construction of which is also motivated by recognition that through-Delta transfers will be compromised if northern California experiences a serious seismic event.

impacts to the economy, this subsidy distorts the value of water. It assigns water an additional value when used to grow crops. Ignoring the larger issue of the benefits/costs of farm subsidies to begin with, attaching the subsidy directly to water use, versus crop acreage, would remove this disincentive for temporary water transfers.

The third is actually a consideration regarding contract structuring consideration – one that is beyond the scope of the model developed in Chapter 3. In general, there is a need to address concerns of the farming community - both through policy actions and contract restructuring - regarding the long-term community-wide impacts of temporary water transfers. These effects, often termed "ripple effects," refer to the impacts of water transfers on third parties, i.e., those parties not directly included in the contract. In farming communities, these third party impacts include revenue loss by those providing auxiliary services to farmers, i.e., mills, crop marketers, and equipment salesmen, and job loss by farm workers. These impacts are what economists call *fiduciary impacts* of restructuring and argue should be ignored on the grounds that they are trumped by the new economic efficiency attained under the restructuring. Job migration and redistribution of incomes occurs across sectors. However, such an argument doesn't apply when the restructuring is temporary. The farmers interact with the millers, the marketers, and the salesmen year after year. Selling water and idling land one year for a handsome price, and at the same time cutting the rest of the dependents out of the value chain, but then resuming "business as usual" in wet years understandably sours business relations.

There is the additional consideration that dips in business during years land is fallowed could potentially destabilize the local economy. If a rice mill sees a halving of its load in years rice land is fallowed, with no financial compensation during these periods, it may not remain economically viable. Milling and other operations do experience some natural fluctuation in annual volume due to favorable/unfavorable growing conditions. The concern is that if large land acreages were to be fallowed in a given year that this effect could be much more pronounced. In order to remain in business, a mill could raise its rate during high-volume years, thereby covering the maintenance/operational costs of the off-years. In practice, there may be several impediments to this approach. First, it may entail charging differentiated rates to all customers based on heterogeneous land fallowing practices. Second, it could create cash-flow problems for the business: multiple dry years with large fallowing tracts could bankrupt the business before it has a chance to "charge back" the operations/maintenance fees in on-years. The current volume of rice production statewide has remained relatively stable at between 44-45 million cwt. There is a healthy amount of competition between mills and marketers in California, which results in some fluctuation in annual volumes processed

by individual facilities. There is also generally reported over-capacity of statewide rice mills, with an estimated capacity to handle up to 80 million cwt. To the extent that there is over-capacity and existing fluctuations in volumes processed, businesses can be expected to adjust to changes in volumes of rice production due to temporary water transfers. Nonetheless, in view of the fairness issues discussed in detail above and the possibility of rice land fallowing to make a dent in the statewide volume of rice production, a contractual solution should be considered.

A contractual solution might be the introduction of retainers. Farmers fallowing land would pay members of the value chain a standard retainer. The fee could be paid annually or quarterly and ensures that the businesses integral to the value chain remain open and at the same time mitigates the community-wide impacts by keeping incomes stable. Retainers for farm employees may be more difficult to negotiate but would also preserve good relations and keep a skilled work force on-hand. Clearly the scope and implementation of these retainers is a complex issue and bears further investigation.

## 5.3 Areas for Future Work

### 5.3.1 Inventory-Price Control

The price-inventory control model developed in Chapter 2 has a number of limitations, which present avenues for future research. To explore these limitations more fully, it is helpful to revisit the basic model structure and each of the underlying assumptions. We recall that the model is formulated as a finite horizon problem, with  $T$  time periods. This implies that there is an end period – a period in which the intermediary's operations will cease. One might object to this assumption on the grounds that the water intermediary will be operating indefinitely, yet the fact that the time horizon can be extended to any large  $T$  and the fact that with even a nominal discount rate the operations far into the future will have a negligible impact on decision-making today make this an unrestrictive assumption. In fact, as discussed below, a model with differentiated time periods lends itself to the more detailed investigation that we would like to do.

Our assumption regarding the *length* of the time period, however, requires more careful consideration. We assumed a yearly time step which was, we argued, in keeping with MWD's annual price review. The implication is, of course, that the price and order decisions are both made once, and only once, per annum. That is, MWD sets a single price and places an

order (or separate orders totalling the optimal order quantity) from its contracts and then does nothing for a year! We have already observed that MWD's current practice is to set prices annually, but what about ordering from contracts? MWD holds a number of different supply contracts, including recently the option contracts discussed in Chapters 3 and 4, and adjusts orders from these contracts in response to demand and available storage.<sup>2</sup> Therefore, annual review of the order decision is not realistic. A potentially messy respecification of the model might shorten the length of each period (resulting in a larger total number of periods) and allow price to be controlled every  $n^{\text{th}}$  period. For instance, if we believe that seasonal resolution suffices regarding ordering decisions, e.g., MWD orders from its various contracts approximately quarterly, then we can allow price to be reviewed every fourth period and ordering policy to be adjusted each period.

The shortening of the time periods, possibly to quarterly durations, leads us to review another assumption: that of stationarity. The assumption that the distribution of supply and demand realizations doesn't change from year to year is a relatively strong assumption. There are larger climatic patterns that impact supply distributions, including the El Nino/La Nina effect and the Pacific Decadal Oscillation (PDO), and, in particular, make the coupling of wet or dry years, respectively, more likely. Demand may also be shifting, depending in large part on the degree to which we believe conservation will offset population increases. These objections notwithstanding, the assumption of stationary demand at the annual time-step is arguably an acceptable approximation, assuming we specify a distribution in accordance with the dominant climatic patterns of the time. At the seasonal time-step, stationarity goes out the window.

It is not only the joint distribution of supply and demand that may vary from period to period. The other parameters of the model may realistically differ from period to period, including the shortage cost and the per-unit delivery cost. These parameters actually also vary within periods – another limitation of the model! We have assumed that these costs can be represented as constant per-unit costs. As already discussed the intermediary holds different contracts, with different costs associated with each, and, hence the cost function will depend on the contract mix used to meet the order-up-to-level (and will not be constant).

The subject of costs bears further consideration – in particular, the per unit cost of water supply, as represented by the order/delivery cost. The approach in our application in Chapter 2 was to approximate this cost based on costs reported, by category, in MWD's annual budget (MWD 2004), as the total reported variable costs divided by the quantity delivered

---

<sup>2</sup>Again, using the term *orders* loosely, in that MWD may be set to receive a certain amount of water, say from SWP, and may be able to refuse the water, if storage is full, or sell or trade it.

that year. This treatment of costs circumvents the equivocal issue of environmental cost associated with water supply.<sup>3</sup> The valuation of ecosystem services and the protection of wildlife is a complicated, yet salient and important, topic.

In closing, we return to our point of departure, the annual review of the pricing decision. We have briefly discussed how we might adapt the dynamic model to account for the fact that prices are adjusted annually, while ordering and supply management decisions are made on a more frequent basis. We might instead consider a policy change, wherein MWD adopts dynamic pricing over a shorter time horizon, possibly in keeping with ordering decisions, e.g., on a quarterly basis. There may be institutional resistance to such a change in policy. However, given the associated efficiency gains, the policy is worth consideration. The value of dynamic pricing over shorter time horizons, or total efficiency gain, should be a question of interest to water intermediaries, which must be considering innovative ways to manage increasingly constrained systems.

### 5.3.2 Contracting

While the topic of inventory-price control might be reasonably considered within the confines of a single intermediary's management policy, the topic of contracting is rightly considered in a much broader social context. Values, such as fairness, and social convention influence, if not entirely determine, what types of contracts are preferable and, ultimately, feasible. When value systems or social convention are challenged, institutional resistance results. As we observed in Chapter 1, the stalled development in the water market is at least in part attributable to a failure to account for resistance to permanent water transfers, which threaten the viability of agricultural communities. It also makes option contracts more appealing than negotiated permanent sales simply on feasibility grounds alone, leaving efficiency out of the picture.

Future work on contracting in the water sector might consider both of these issues in more detail. First, the issue of feasibility, or viable mechanisms for water allocation. The existing literature discusses institutions and conventions in the water sector in detail. It does not, however, take the next step and consider, given these existing institutional structures, which types of contracts will be feasible. Our discussion of bilateral option contracts as a means of reducing institutional resistance to water reallocation might be considered one example of this, but there are many other contracting possibilities – and a number of different in-

---

<sup>3</sup>MWD does invest in a number of fish protection and conservation programs, so the environmental costs, or at least the cost of remediation, are not entirely absent from our estimate.

stitutional settings outside California. A simple extension to the bilateral framework might include consideration of the feasibility of long-term, or multi-year, option contracts. Future investigation in this area seems particularly important in light of an observation made by Barton H. Thompson, that contracts that are set in the early phases of the water market's development will in all likelihood serve as templates for future contracting.

The second question, that of efficiency, abstracts from institutional constraints and asks, *what is the socially optimal mix of temporary and permanent transfers under uncertainty?* To be more concrete, we return to a two-period world. In the first period, the social planner decides how to assign (or reassign) property rights, e.g., the permanent transfers, and sets the terms for the temporary transfers. An unimpeded social planner, interested in welfare maximization, would implement a flexible transfer plan that in effect allows him to postpone the decision to the second period when the uncertainty realizes. These option contracts would allocate the resource to the user with the highest value in the second period. That is, all consumers would have an option to buy the good, with the execution price indexed to the market-clearing price in the second period. Hence, the corresponding optimal contracting mix is all temporary transfers, or option contracts.<sup>4</sup>

However, what if the social planner is impeded, in the sense that he faces constraints on the admissible second-period transfers. For instance, suppose that there is a probability of non-delivery attached to the temporary transfer. That is, all transfers that are executed in the second period will not go through with some positive probability. In this simple model, it no longer suffices to simply endow all consumers with options indexed to the market-clearing price. The social planner may wish to ensure a consumer with a higher expected value for the good at least some piece of the pie. Hence a permanent transfer is necessary. The question is, how much should be allocated through this permanent transfer? The answer must depend on the both the potential future loss to the customer from not having the good and the probability of non-delivery associated with the transfer contract, where the contract could be simply an option to buy the good on the market at an indexed price. The formal characterization of such conditions is a future research project.

Both the aforementioned areas for future work expand the scope of Chapters 3 and 4. There are a number of limitations associated with the models therein, also areas for future investigation. Again, it is helpful to revisit the basic model structure and consider the underlying assumptions. We recall that the bilateral option contracting model in Chapter 3 has the following general structure: a single buyer and seller contract with perfect information, where the seller holds the price-setting power. The two uncertainties are (1) the second-period

---

<sup>4</sup>This is the socially optimal allocation discussed briefly in Chapter 3 in the context of two players.

(excess) demand that the buyer will face and (2) the expected spot market price that, alternately, the seller or both the seller and the buyer will face. We also introduce a third uncertainty in our general exposition, which features in the application of our model: the probability of an infrastructure failure, e.g., the probability of non-delivery.

The assignment of the price-setting power to the seller is a somewhat restrictive assumption in light of current contracting in the water market, which is perhaps more accurately described as a negotiation. That said, in a recent meeting with San Diego County Water Authority (SDCWA), which has entered into two option contracts this year, it became apparent that the farmers are in the stronger negotiation position. SDCWA needs the water more than the farmers need to sell it. Furthermore, in the advent of a more developed market with potential involvement by third parties, the model will be generally applicable: sellers do set prices! The contracting model for capacity in the electricity market from Wu et al. (2002) was developed under similar motivations, where sellers set contract prices.

The three uncertainties in the model accurately reflect the conditions in the contracting market, with the inclusion of the delivery uncertainty an addition to the model owing to discussions with SDCWA. The treatment of the two uncertainties as independent can reasonably be called into question when there is a spot market with joint access, given the positive correlation between demand and dry water years discussed in Chapter 2. An extension of the model to handle correlated uncertainty when there is joint spot market access would be worthwhile.

The discussion in Chapter 4 opens the door to extensions in several directions. We provide an analysis of seller competition, with complete information, under a simplified demand structure. It would be interesting to extend this analysis to a more realistic demand structure, although, as noted, this requires the treatment of two prices. Our numerical study demonstrates that, as in the single-price case, there are under-capacity regions where there are pure strategy equilibria – and regions where we fail to find such pure strategy equilibria. The question of interest in such a study, with regard to the water market, is whether we get more competitive prices under seller competition, e.g., whether it improves efficiency. Generally, we would expect this to be the case. We see, however, that when there are predictable pure strategy Nash equilibria, it is in the under-capacity regions where prices are high, calling this assertion into question.

As we note in Chapter 4, the introduction of many sellers makes the assumption of complete information regarding sellers' capacities less robust. We describe briefly the possibility of a multi-unit auction, which would be efficient but where clearly the distributional outcomes



depend on the number of sellers participating. Hence, the attractiveness of the auction to the buyer, presumably conducting the auction, may be quite sensitive to his estimate of the number of participants. This raises a number of questions regarding viability of multi-unit auctions, e.g., when are they appealing to buyers and when are they not? A more extensive investigation is warranted, one that may additionally relax the assumption of complete information regarding sellers values.

In Chapter 1, we touched on the various aspects of the water intermediary's decision problem summarized in Figure 1. We have addressed just three areas of interest – rate-setting, integrated with ordering from contracts, or procurement, and contracting. The other six areas – long-term supply contracts, capacity expansion, conservation programs, reservoir/storage operations, regulatory influence, and rationing policy – have remained unexplored in this thesis. To the extent that these other areas could be integrated into a common decision-framework, would be interesting to explore. Although, this suggests more of a large-scale system model approach, and there is certainly some work in this area (see, e.g., Lund 2001). Treatment of each topic individually is interesting in its own right, and there is a literature associated with each of these other topics. For instance, there is an extensive water management literature treating reservoir/storage operations. The capacity expansion decision is one treated in the peak-load pricing literature, integrated with the tariff decision. It would be interesting to explicitly couple this decision with the joint inventory-price control model. Different rationing schemes are discussed in the economics literature, including random rationing versus rationing on a willingness-to-pay basis (efficient rationing). This is also discussed in the peak-load pricing literature in conjunction with shortage costs (which clearly depend on the rationing scheme in place). MWD's rationing scheme, which was instituted during California's drought in the early 1990s, is discussed in Chapter 2. A review of this policy in terms of efficiency implications might be a starting point for a more comprehensive review of rationing, which may not be an uncommon phenomena in the future.

# Bibliography

- [1] ALLEN, B., MARTIN, H. (1993) "Bertrand-Edgeworth Duopoly with Proportional Residual Demand," *International Economic Review*, Vol. 34, No. 1, pp. 39–60.
- [2] AMERICAN WATER WORKS ASSOCIATION (AWWA) (2000) "Principles of Water Rates, Fees, and Charges, Fifth Edition," AWWA, Denver, CO.
- [3] ANDRIYCHENKO, O., GIRNIUS, A., SAHA, A. (2006) "Complementary Goods: Prices and Consumer Welfare Under Duopoly and Monopoly," *International Journal of the Economics of Business*, Vol. 13, No. 3, pp. 373–386.
- [4] ARROW, K., HARRIS, T., MARSCHAK, J. (1951) "Optimal Inventory Policy," *Econometrica*, Vol. 19, No. 3, pp. 250–272.
- [5] ARELLANO, M.S., SERRA P. (2007a) "A Model of Market Power in Electricity Industries Subject to Peak Load Pricing," *Energy Policy*, Vol. 35, No. 10, pp. 5130–5135.
- [6] ARELLANO, M.S., SERRA, P. (2007b) "Spatial Peak-Load Pricing," *Energy Economics*, Vol. 29, No. 2, pp. 228–239.
- [7] ATHEY, S. (2002) "Monotone Comparative Statics Under Uncertainty," *Quarterly Journal of Economics*, Vol. 117, No. 1, pp. 187–223.
- [8] BENOIT, J., KRISHNA, V. (1987) "Dynamic Duopoly: Prices and Quantities," *Review of Economic Studies*, Vol. 54, No. 177, pp. 23–35.
- [9] BERTRAND, J. (1883) "Review of 'Theorie Mathematique de la Richesse Sociale' and 'Recherches sur les Principes Mathematiques de la Theorie des Richesses,'" *Journal des Savants*, pp. 499–508.
- [10] BOITEUX, M. (1956) "Sur la Gestion des Monopoles Publics Astreints a l'Equilibre Budgetaire," *Econometrica*, Vol. 24, pp. 22–44; English translation: (1971) "On the

- Management of Public Monopolies Subject to Budgetary Constraints," *Journal of Economic Theory*, Vol. 3, pp.219–240.
- [11] BÖS, D. (1985) "Public Sector Pricing," in: Auerbach, A.J., Feldstein, M. (Eds.), *Handbook of Public Economics, Vol. I*, Elsevier, Amsterdam, NL, pp. 128–211.
- [12] CAYSEELE, P., FURTH, D. (1996) "Bertrand-Edgeworth Duopoly with Buyouts or First Refusal Contracts," *Games and Economic Behavior*, Vol. 16, No. 2, pp. 153–180.
- [13] CHEN, F.Y., RAY, S., SONG, Y. (2006) "Optimal Pricing and Inventory Control Policy in Periodic-Review Systems with Fixed Ordering Cost and Lost Sales," *Naval Research Logistics*, Vol. 53, No. 2, pp. 117–136.
- [14] COURNOT, A. (1838) "Recherches sur les Principes Mathematiques de la Theorie des Richesses" Paris: 1838 English translation: (N. Bacon, trans.) "Researches into the Mathematical Prinsiples of the Theory of Wealth," New York: Macmillan & Company, 1897.
- [15] COWAN S. (2003) "Optimal Risk Allocation for Regulated Monopolies," *Journal of Public Economics*, Vol. 88, No. 1-2, pp. 285–303.
- [16] CREW, M.A., FERNANDO, C., KLEINDORFER, P. (1995) "The Theory of Peak Load Pricing: A Survey," *Journal of Regulatory Economics*, Vol. 8, No. 3, pp. 215–248.
- [17] DE BORGER, B. (2000) "Optimal Two-Part Tariffs in a Model of Discrete Choice," *Journal of Public Economics*, Vol. 76, No. 1, pp. 127–150.
- [18] DOORMAN, G.L. (2005) "Capacity Subscription: Solving the Peak Demand Challenge in Electricity Markets," *IEEE Transactions on Power Systems*, Vol. 20, No. 1, pp. 239–245.
- [19] ECONOMIDES, N., VIARD, B.V. (2006) "Pricing of Complementary Goods and Network Effects," Working Paper, Department of Economics, New York University, New York, NY.
- [20] FEDERGRUEN, A., HECHING A. (1999) "Combined Pricing and Inventory Control Under Uncertainty," *Operations Research*, Vol. 47, No. 3, pp. 454–475.
- [21] FISCHER, R., SERRA P. (2002) "Energy Prices in the Face of Plant Indivisibilities," *Energy Economics*, Vol. 25, No. 4, pp. 303–314.

- [22] GALLEGO G., VAN RYZIN, G. (1994) "Optimal Dynamic Pricing of Inventories with Stochastic Demand over Finite Horizons," *Management Science*, Vol. 40, No. 8, pp. 999–1020.
- [23] GALLEGO G., OZER, O. (2001) "Integrating Replenishment Decisions with Advance Demand Information," *Management Science*, Vol. 47, No. 10, pp. 1344–1360.
- [24] GARCIA, A., REITZES, J.D., STACCHETTI, E. (2001) "Strategic Pricing When Electricity is Storable," *Journal of Regulatory Economics*, Vol. 20, No. 3, pp. 223–247.
- [25] HANAK, E. (2003) "Who Should Be Allowed to Sell Water in California? Third-Party Issues and the Water Market," Public Policy Institute of California, San Francisco, CA.
- [26] HANAK, E. (2005) "Stopping the Drain: Third-Party Responses to California's Water Market," *Contemporary Economic Policy*, Vol. 23, No. 1, pp. 59–77.
- [27] HANAK, E., HOWITT, R.H. (2005) "Incremental Water Market Development: The California Water Sector 1985–2004," *Canadian Water Resources Journal*, Vol. 30, No. 1, pp. 73–82.
- [28] HERTER, K. (2007) "Residential Implementation of Critical-Peak Pricing of Electricity," *Energy Policy*, Vol. 35, No. 4, pp. 2121–2130.
- [29] HIRSCHBERG, J.G. (2000) "Modelling Time of Day Substitution Using the Second Moments of Demand," *Applied Economics*, Vol. 32, No. 8, pp. 979–986.
- [30] JANSSEN, M., RASMUSEN, E. (2002) "Bertrand Competition Under Uncertainty," *The Journal of Industrial Economics*, Vol. 50, No. 1, pp. 11–21.
- [31] KARLIN, S., RUBIN, H. (1965) "The Theory of Decision Procedures for Distributions with Monotone Likelihood Order," *Annals of Mathematical Statistics*, Vol. 27, No. 2, pp. 272–299.
- [32] KARLIN, S., RINOTT, Y. (1980) "Classes of Orderings of Measures and Related Correlation Inequalities. I. Multivariate Totally Positive Distributions," *Journal of Multivariate Analysis*, Vol. 10, No. 4, pp. 467–498.
- [33] KREPS, D.M., SCHEINKMAN, J.A. (1983) "Quantity Precommitment and Bertrand Competition Yield Cournot Outcomes," *Bell Journal of Economics*, Vol. 14, No. 2, pp. 326–337.

- [34] LI, Q., ZHENG, S. (2006) "Joint Inventory Replenishment and Pricing Control for Systems with Uncertain Yield and Demand," *Operations Research*, Vol. 54, No. 4, pp. 696–705.
- [35] LIN, K.Y. (2006) "Dynamic Pricing with Real-Time Demand Learning," *Production, Manufacturing, and Logistics*, Vol. 174, No. 1, pp. 522–538.
- [36] MATSUKAWA, I. (2001) "Household Response to Optional Peak-Load Pricing of Electricity," *Journal of Regulatory Economics*, Vol. 20, No. 3, pp. 249–267.
- [37] MILGROM, P., WEBER, R. (1982) "A Theory of Auctions and Competitive Bidding," *Econometrica*, Vol. 50, No. 5, pp. 1089–1122.
- [38] MILGROM, P., SHANNON, C. (1994) "Monotone Comparative Statics," *Econometrica*, Vol. 62, No. 1, pp. 157–180.
- [39] MILLS, E.S. (1959) "Uncertainty and Price Theory," *Quarterly Journal of Economics*, Vol. 73, No. 1, pp. 116–130.
- [40] NISAN, N., ROUGHGARDEN, T., TARDOS, E., VAZIRANI, V., EDS. (2007) *Algorithmic Game Theory*, Cambridge University Press, Cambridge.
- [41] MILLS, E.S. (1959) "Uncertainty and Price Theory," *Quarterly Journal of Economics*, Vol. 73, No. 1, pp. 116–130.
- [42] NG, Y., WEISSER, M. (1974) "Optimal Pricing with a Budget Constraint – The Case of the Two-part Tariff," *Review of Economic Studies*, Vol. 41, No. 3, pp. 337–345.
- [43] OI, W.Y. (1971) "A Disneyland Dilemma: Two-Part Tariffs for a Mickey Mouse Monopoly," *Quarterly Journal of Economics*, Vol. 85, No. 1, pp. 77–96.
- [44] OSBORNE, M.J., PITCHIK, C. (1986) "Price Competition in a Capacity-Constrained Duopoly," *Journal of Economic Theory*, Vol. 38, No. 2, pp. 238–260.
- [45] PETRUZZI, N.C., DADA, M. (1999) "Pricing and the Newsvendor Problem: A Review with Extensions," *Operations Research*, Vol. 47, No. 2, pp. 183–194.
- [46] PETRUZZI, N.C., DADA, M. (2002) "Dynamic Pricing and Inventory Control with Learning," *Naval Research Logistics*, Vol. 49, No. 3, pp. 303–325.
- [47] PINEAU, P., ZACCOUR, G. (2007) "An Oligopolistic Electricity Market Model with Interdependent Segments," *Energy Journal*, Vol. 28, No. 3, pp. 165–185.

- [48] POLATOGLU, L.H., SAHIN, I. (2000) "Optimal Procurement Policies Under Price-Dependent Demand," *International Journal of Production Economics*, Vol. 65, No. 2, pp. 141–171.
- [49] RAJARAM, K., YIN, R. (2007) "Joint Pricing and Inventory Control with a Markovian Demand Model," *European Journal of Operational Research*, Vol. 182, No. 1, pp. 113–126.
- [50] RAMSEY, F. (1927) "A Contribution to the Theory of Taxation," *Economic Journal*, Vol. 37, No. 145, pp. 47–61.
- [51] REYNOLDS, S., WILSON, B. (2000) "Bertrand-Edgeworth Competition, Demand Uncertainty, and Asymmetric Outcomes," *Journal of Economic Theory*, Vol. 92, No. 1, pp. 122–141.
- [52] SKIERA, B., SPANN, M. (1999) "The Ability to Compensate for Suboptimal Capacity Decisions by Optimal Prices," *European Journal of Operational Research*, Vol. 118, No. 3, pp. 450–463.
- [53] SMITH, J., MCCARDLE, K. (2002) "Structural Properties of Stochastic Dynamic Programs," *Operations Research*, Vol. 50, No. 5, pp. 796–809.
- [54] TALLURI, K.T., VAN RYZIN, G.J. (2004) *Theory and Practice of Revenue Management*, Springer: New York, NY.
- [55] THOMAS, L.J. (1974) "Price and Production Decisions with Stochastic Demand," *Operations Research*, Vol. 22, No. 3, pp. 531–518.
- [56] THOMPSON, B.H. (1993) "Institutional Perspectives on Water Policy and Markets," *California Law Review*, Vol. 81, No. 3, pp. 673–764.
- [57] THOWSEN, G.T. (1975) "A Dynamic Non-Stationary Inventory Problem for a Price/Quantity Setting Firm," *Naval Research Logistics Quarterly*, Vol. 22, No. 3, pp. 461–476.
- [58] TIROLE, J. (1988) *The Theory of Industrial Organization*, MIT Press, Cambridge, MA.
- [59] TOMKINS, C.D., WEBER, T.A. (2008) "Option Contracting in the California Water Market," Working Paper, Department of Management Science and Engineering, Stanford University, Stanford, CA.

- [60] Weber, T.A. (2007) "Capacity-Constrained Bertrand Competition with Asymmetric Information," Unpublished Manuscript, Department of Management Science and Engineering, Stanford University, Stanford, CA.
- [61] WEITZMAN, M.L. (1974) "Prices vs. Quantities," *Review of Economic Studies*, Vol. 41, No. 4, pp. 477–491.
- [62] WHITIN T.M. (1953) *The Theory of Inventory Management*, Princeton University Press, Princeton, NJ.
- [63] WHITIN T.M. (1955) "Inventory Control and Price Theory," *Management Science*, Vol. 2, No. 1, pp. 61–68.
- [64] WU, D.J., KLEINDORFER, P.R., ZHENG, J.E. (2002) "Optimal Bidding and Contracting Strategies for Capital-Intensive Goods," *European Journal of Operational Research*, Vol. 137, No. 3, pp. 657–676.
- [65] ZABEL, E. (1972) "Multiperiod Monopoly Under Uncertainty," *Journal of Economic Theory*, Vol. 5, No. 3, pp. 524–536.