INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

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 bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI 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.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA

UMI®
800-521-0600

RICE UNIVERSITY

Inter-Temporal and Spatial Choice Between Ground and Surface Water in Municipalities With Land Subsidence.

by

Sylvestre Gaudin Hultberg

A Thesis Submitted
in Partial Fulfillment of the
Requirements for the Degree
Doctor of Philosophy

Approved, Thesis Committee

Peter Mieszkowski, Chairman
Professor, Economics

Peter Hartley
Professor, Economics

Robin Sickles
Professor, Economics

Mark Wiesner
Professor, Environmental Sciences and Engineering

Houston, TX

June 1999

UMI Number: 9969271

UMI Microform 9969271
Copyright 2000 by Bell & Howell Information and Learning Company.
All rights reserved. This microform edition is protected against
unauthorized copying under Title 17, United States Code.

Bell & Howell Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

ABSTRACT

Inter-Temporal and Spatial Choice Between Ground and Surface Water in Municipalities With Land Subsidence.

by

Sylvestre Gaudin Hultberg

The impact of land subsidence on optimal inter-temporal use of groundwater is analyzed for a growing city where surface water is a feasible alternative and the goal is to maximize net social benefits. The problem is dynamic, since pumping and subsidence externalities increase costs in all future periods when groundwater is mined. It is also spatial because surface water costs increase with distance. The distance to which surface water is delivered is shown to depend on population densities, city size, and the proportion of surface water to groundwater use.

The model is solved using dynamic programming and numerical techniques. Parameter values are chosen to simulate the case of Houston. The demand side is specified using consumer surplus from a Stone-Geary demand function, estimated with panel data. The simulation results confirm the analytical predictions that at first only groundwater is used and as demand grows, water levels go down and subsidence occurs. When the social marginal cost of groundwater reaches the marginal cost of surface water, it becomes optimal to use surface water. Because of increasing returns to scale in water treatment, surface water

is used in large quantities right away. A period follows during which water levels increase and subsidence stops. Water levels start decreasing again when decreasing returns to dispersion dominate. Unless a water treatment plant is built in a different location, the use of groundwater keeps increasing and subsidence eventually resumes.

The simulations indicate that if surface and groundwater have been used optimally in Houston, the discount rate must be low or subsidence costs are assumed very high. The recent reversal in water level declines is explained by increasing returns in treatment but because of increasing dispersion water levels should be allowed to drop again, as long as new treatment plants are built in different locations before subsidence resumes.

ACKNOWLEDGEMENTS

Many individuals were at one time or another involved in this long process that required many different skills and lots of information. First and foremost I wish to thank my thesis committee for their support along the way to finishing this task that often felt overwhelming. I could not have found a better committee. The complementarity of their skills and emphasis shaped my work. My deepest appreciation for my thesis director, Peter Mieszkowski, who put up with all my emotional swings, especially at the early stages, and constantly reminded me about meaning, consistency, and impact, seeing always the "big picture". I appreciated the great deal of economic intuition that I was privileged to share for the last few years. My thanks to Robin Sickles, who came into the committee at a time when I had little confidence in the feasibility of this research. He helped me focus on a manageable econometrics problem, and returned my confidence through constant positive reinforcement and genuine belief that I was up to the task. I also benefited greatly from his wide knowledge in econometrics. I am indebted to Peter Hartley who methodologically guided me through the technical aspects and logical consistency of the dynamic programming model. From the first year in the Ph.D. program he made me believe that there is a way to combine techniques and economic insight. Mark Wiesner and Phil Bedient from the Environmental Engineering helped me at the early stages to get through the non-economic aspects of my research that proved crucial in the statement of my problem. I appreciated their time and expertise. I also appreciated the collaboration of Ron Griffin, from the department of agricultural economics at Texas A&M, who provided me with the panel data I use in chapter 3 and spent time discussing the more general picture of my research. Tom Edgar from the department of Civil Engineering at the University of Wyoming was very helpful and I greatly appreciated his time.

In addition, I wish to thank the Energy and Environmental Systems Institute of Rice University and its sponsors for providing initial funding for this research through their Research Initiation Grant program. Without this grant, I would be discussing Brazilian income distribution issues! I also thank my fellow graduating friend, Weiyu Gao, for his comments and suggestions on the econometrics chapter, and for sharing his work that also involves dynamic programming. Many individuals from the community were involved in providing me information and expertise on water issues and data. In particular, I wish to thank the following people who met with me in 1997 and 1998, some for extended periods of time: Stephen Densmore, Roger Quincy, and Norman Alford from the TWDB in Austin; Bob Gabrysch from the HGSD; Bill Healer from the City of Deer Park; Ron Welch from Municipal Information Services, and Chuck Settle from the City of Houston, Public Works and Engineering.

As the length of these acknowledgments testify, this was a long journey. Along the way, I married Patrik. I cannot express the role of my wonderful and beloved husband in this endeavor. We shall remember our intense brainstorming sessions and the last stretch of sleepless nights for some time. Thank you Patrik, and this is also for you.

TABLE OF CONTENTS

CHAPTER 1 INTRODUCTION 1

1.1 Motivation and objectives 1

1.2 Relevance 3

1.3 The Theoretical Problem 5

1.4 Place in the Economic Literature 9

1.5 Outline of the Thesis 12

CHAPTER 2 THE THEORETICAL MODEL: METHODOLOGY AND FINDINGS 16

2.1 Optimal Spatial Allocation of Surface and Groundwater Use 16

2.2 The Mathematical Model: Dynamic Programming Approach to the Optimal Control Social Planning Problem. 21

2.2.1 Model Variables 22

State Variables 22

Control Variables: 23

Other Variables: 24

2.2.2 Model Functional Equations 24

Transition Equation for the State Variables: 24

Gross Benefit Function: 25

Cost Functions : Groundwater 26

Cost functions: Surface Water 27

2.2.3 The Dynamic Optimization Problem 28

Objective Function 28

Dynamic Programming Formulation 29

2.3 Expected Results: Behavior of State Variables and Controls 32

Base Model: No Further Investment in Surface Water Treatment 32

Extension to the Case of Lumpy Investment in Surface Water Treatment. 35

CHAPTER 3 CHOOSING DEMAND PARAMETERS: EVALUATION OF THE STONE-GEARY FORM USING PANEL DATA ESTIMATION. 37

3.1 Introduction 37

3.2 Data and Functional Forms 39
3.2.1 Data 39
3.2.2 Functional Forms 42
The Generalized Cobb-Douglas (GCD) Demand Function 42
The Stone-Geary Demand Function 43

3.3 Generalized Cobb-Douglas Estimation 46
3.3.1 Estimation Methods 46
3.3.2 GCD Results 47

3.4 Stone-Geary estimation 53
3.4.1 Estimated Equations 54
3.4.2 Transformation of Variables for Nonlinear GLS 55

3.5 Stone-Geary Results and Comparison to GCD 56

3.6 Conclusion 62

CHAPTER 4 NUMERICAL SIMULATIONS AND RESULTS 68

4.1 Detailed Functional Forms and Parameterization Using the Case of Houston, Texas 69
4.1.1 Transition Equations and Bounds of State Variables 69
Change in water levels 69
Changes in the minimum historic head: 71
Bounds for H and M : 72
Population growth and bounds: 73
4.1.2 Benefits: Gross Consumer Surplus 74
4.1.3 Groundwater Costs 76
Groundwater Pumping Costs 76
Groundwater Subsidence Costs 78
4.1.4 Surface Water Costs 80
Surface Water Treatment Costs 80
Surface Water Transmission Costs 82
4.1.5 Spatial Parameters 84
4.1.6 Discount Factor and Summary 87

4.2 Simulation 87
4.2.1 Methodology 87
4.2.2 Results 94

4.3 Sensitivity analysis 103
4.3.1 Sensitivity to the discount factor β 103
4.3.2 Sensitivity to subsidence cost parameters (values of $\overline{c}$) 108
No subsidence costs ($\overline{c} = 0$) 108

High subsidence costs ($\bar{c} = 300$ and $\bar{c} = 400$) 111
4.3.3 Sensitivity to the spatial parameters (l, λ, K) 112

CHAPTER 5 A CASE STUDY OF THE HOUSTON-GALVESTON AREA, TEXAS 120

5.1 The Facts 120

5.2 Model Predictions and the Case of Houston-Galveston. 127

5.3 Regulation Against Subsidence 129
5.3.1 The Harris-Galveston Coastal Subsidence District 129
5.3.2 The District Plans 131

5.4 Analysis of the HGSD Regulation-Spatial Issues 133
5.4.1 The 1992 Changes are Consistent with a Better Cost-Minimizing Path 134
5.4.2 Case against Equal Use of Groundwater: the Tradable Permit Alternative. 135

GENERAL SUMMARY AND CONCLUSIONS 138

REFERENCES 143

APPENDIX 150

CHAPTER 1

Introduction

1. Motivation and objectives

Managing water resources and planning for future water supply is a complex task in growing cities where both groundwater and surface water are potentially available. Although politics plays a role, choices are mostly driven by economic considerations. Looking only at direct private costs of procuring potable water, groundwater is usually a first choice for municipal water supply because the amount of treatment required is minimal, while surface water needs extensive treatment. Because surface water has become more contaminated at the same time that drinking standards have become more stringent, surface water has become more costly to treat. Moreover, surface water has become more scarce as rapid urban growth has caused large transfers of water use from agriculture to municipalities.[1]

Having mostly drawn on groundwater reserves in the past, many municipalities have rapidly depleted their groundwater reserves. The term "groundwater mining" is used when the amount of groundwater withdrawn exceeds the amount of recharge, so groundwater levels decrease. Groundwater mining is accompanied by increased pumping costs, water quality degradation, and, sometimes land subsidence. When land subsidence occurs, it constitutes the largest external cost associated with the mining of aquifers. Land subsidence is a scientifically well documented phenomenon associated for a large part to excessive groundwater withdrawals (Waltham, 1984; Carr et al., 1985). As water is extracted in excess of

[1] For example, the Texas Water Development Board estimates that in the year 2040, municipalities will be the main consumers of water in Texas.

recharge –the aquifer is being mined– pressure in the confined aquifer falls, causing water to migrate from clays to sands. Clay pores are now empty and the weight of overlying layers causes the clays to compact, which leads to land subsidence.[2] There are large costs associated with land subsidence. They are mostly due to increased flooding, tidal encroachment into coastal areas, and structural damage to infrastructure and buildings. Subsidence can be stopped but is mostly irreversible. The clays have very low elasticity and there is very little rebound.[3] Letting subsidence occur can cause considerable future costs that accumulate even after subsidence has ceased. Indeed, if the costs in terms of structural damage disappear when land subsidence stops, the costs of increased flooding remain, and some land becomes permanently damaged or simply lost for most economic uses. Acting too late can render the groundwater source unusable, as storage capacity declines and contamination becomes more probable. The implications are serious, since groundwater loses its option value as a buffer against reduced supply of surface water during periods of drought (Tsur and Issar, 1989; Tsur and Zemel, 1995). Because these costs are not directly perceived by groundwater providers, there is an incentive to use too much groundwater. The social cost of using groundwater greatly increases when groundwater is mined, especially when groundwater mining creates land subsidence. It has been argued that, if the external costs of land subsidence were included, groundwater would be much more expensive than surface water, usually perceived as the expensive alternative (Jones and Larson (1975), Teutsch, 1978).

Should the cities with land subsidence problems completely switch to surface water? How much should groundwater be mined before surface water is used? Should the groundwater source be fully depleted before alternative sources are used? Should groundwater and surface water be used simultaneously? Many cities in the US and across the world are faced with such problems, and the sustainability of economic development in these cities depends

[2] There is evidence that the level of the land surface decreases linearly with the log of the effective stress applied to the formation, effective stress being actual stress (caused by the weight of soils above the aquifer) minus water pressure. Stress remains constant but water pressure decreases as water levels decrease, causing subsidence.

[3] Average rebound is of the order of 5 percent, ranging from about 1 to 10 percent depending on the geological characteristics of the aquifer.

largely on improved long-run planning and management of scarce water resources. Our goal is to understand the determinants of optimal choice between groundwater and surface water for municipal water supply over time. The problem has an important spatial element that needs to be addressed: when different parts of the city draw from the same aquifer for groundwater and potentially from the same reservoir and treatment plant for surface water, they are likely to face different surface water costs depending on their distance to the treatment plant. This should impact the optimal spatial allocation of groundwater within the metropolitan area in order to minimize costs. First, we need to solve the spatial problem to understand how the city planner should allocate groundwater usage among the different areas of the city. Then, given this best spatial allocation, we need to solve for the optimal amount of water from each source every period, given the social costs of groundwater and surface water provision. This will determine the optimal share of groundwater in total water usage, the amount of groundwater depletion, and the cumulative amount of subsidence.

2. Relevance

There are many well know examples of groundwater depletion and subsidence problems caused by rapidly increasing population. They include, among many others, the cases of Bangkok, Mexico City, Venice, and, in the US, some parts of California (Santa Clara and San Joaquin Valleys), Florida, and Texas.[4] These cities were exclusively using groundwater before they became aware of external mining costs. They are now all using part or all surface water, or are at least studying potential alternative sources of water supply.

The city of Bangkok has probably the highest current rate of uncontrolled subsidence. The City has both surface and ground water available, but surface water must be imported from about 40 miles upstream mostly because of water quality problems in the Chao Phraya

[4] See Waltham (1984) on regional subsidence due to fluid withdrawal (pp. 117-130), Nickum and Easter (1994) on cases in Asia and the Pacific, especially the case of Bangkok. See also the USGS Internet site on subsidence in the United States at http://water.wr.usgs.gov/subsidence/ls_3.html.

River. In addition, upstream rice irrigators and fruit growers have conflicting claims to the available surface water. In 1983 a decree directed the Metropolitan Water Authority (MWA) to switch to surface based water systems. However, implementation has been slow and the MWA has not been able to keep pace with the rapid growth of the city's population and economy. Moreover, non-enforcement or low penalties often render water laws quite ineffective. The MWA does not have control over all the water withdrawn because many illegal wells are used and because metering is poor. Land subsidence continues and causes increasing damage as more is mined from the aquifer (Nickum and Easter, 1994).

Mexico City addressed the problems of groundwater depletion and subsidence after considerable damage had already occurred, some permanent and irreversible. Subsidence in Mexico City has reached 30 feet. Some well casings are now protruding 15 feet above street level. Groundwater use from the Mexico Valley aquifer has been restricted since the early 1980's in response to the lowering of the water table, land subsidence, and the deterioration of water quality. While subsidence rates have declined, the city must now rely on surface water imported from the Cutzamala river, which must be pumped a vertical distance of one thousand meters, using a 180 kilometer pipeline. The average incremental cost of water, excluding treatment and distribution, rose 55 percent in about ten years (World Bank, 1993). Aquifer water is no longer a source of water for the city, so in case of severe droughts, widespread shortages are likely.

Between 1908 and 1969 Venice sank 220 millimeters below sea level. After 1925, the rate of subsidence became far greater than the earlier natural subsidence rate. The main cause of this accelerated subsidence was groundwater extraction in excess of recharge. The frequency of spring high tide increased from once per decade in 1900 to nearly five times a year in the 1980's. Severe pumping restrictions reversed the head decline in 1969. These allowed a very limited rebound in land surface, and eventually reduced the rate of subsi-

dence to its low natural unstoppable rate. The costs associated with the periodic flooding of Venice due to accumulated subsidence are large and continue to accumulate (Waltham, 1984).

In the United States, annual costs of subsidence mitigation have been estimated to be more than $30 million for Santa Clara County, California (Silicon Valley); $10 million for the state of Florida; $30 million for the city of New Orleans; and over $180 million for the San Joaquin Valley, California.[5] A major area in the United States with similar problems is the Houston-Galveston area in Texas. We will discuss the Houston-Galveston case in detail in the last chapter.

3. The Theoretical Problem

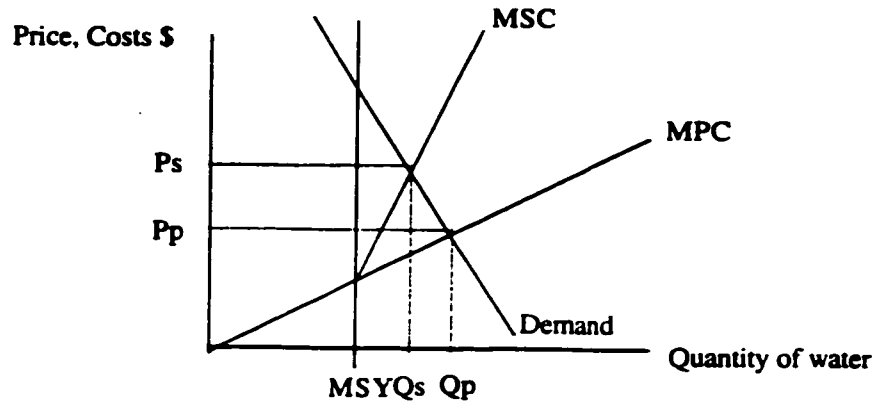
Long term planning in groundwater use is important due to the fact that groundwater is a renewable resource which may be depleted if the rate of use is permanently greater than the rate of recharge, a situation usually referred to as aquifer mining. There are four categories of external costs associated with groundwater mining. These are i) increased pumping costs, ii) costs associated with the deterioration of water quality (including salt water encroachment), iii) costs associated with land subsidence and, when surface and ground water are interconnected, iv) costs associated with the reduction of surface water flows. We concentrate here on the costs imposed by subsidence, which have largely been ignored in the economic literature. However, our analysis will also incorporate the effects of aquifer mining on pumping costs.

In deciding whether to use ground or surface water, municipal districts (or private individuals or firms) look at the direct costs of providing drinking water from each source. When there are several municipal entities sharing the same aquifer, the cost of subsidence is not directly perceived by the utility districts. The usual externality problem occurs, lead-

[5] Numbers from the USGS Internet site at http://water.wr.usgs.gov/subsidence/ls_3.html.

ing to overuse of the resource. In figure 1, MPC represents the marginal private costs of pumping groundwater. MSC represents the marginal social cost. MSC = MPC + marginal external cost. Up to the safe yield (Maximum Sustainable Yield or MSY), there are no external costs associated with the extraction of groundwater since water levels do not change. Beyond the safe yield, external costs start to accumulate. Qp represents the amount of groundwater extracted given the private cost of supplying groundwater, while Qs represent the socially optimal level of pumping. Ps represents the price at which groundwater should be sold to reflect all social costs.

Figure 1: Overproduction of Groundwater with Negative Mining Externality



Since the use of groundwater in one period affects costs in subsequent periods, the problem is dynamic in nature. However, the simple externality analysis above is enough to show that only a coordinated use of an aquifer, possibly by a social planner in charge of deciding on the overall use of the resource, can be socially optimal.[6] This is the approach we take.

What could be the economic forces encouraging the conservation of groundwater? Three possibilities come to mind. First, changes in the cost of groundwater relative to the alterna-

[6] An alternative approach may involve defining, allocating, and enforcing property rights to the available water, while allowing individuals adversely affected by subsidence to sue for damages. Even so, equal restrictions in the total quantity of water rights would probably be needed to achieve an efficient outcome.

tive sources will probably end aquifer mining before the stock of groundwater is depleted. As the cost of groundwater increases, alternative sources become more attractive. Since the resource is renewable, a reduction of pumping rates below rates of recharge can reduce these costs or maintain them at a level comparable to surface water costs. Some aquifer mining may resume to reach a new steady state if new technologies modify the costs of depletion.[7] However, in this dissertation, we ignore these changes in technology because they would complicate the analysis and are unlikely to affect the cost of subsidence as much as other costs.

The second possible motive for conserving groundwater relates to increasing demand. Rubio et al. (1994) find that if the interest (discount) rate is greater than the rate of growth of water demand, optimal management of an aquifer implies a rate of use permanently greater than the rate of recharge, causing depletion of groundwater reserves up to a minimum "survival" level that needs to be set exogenously on geological considerations. If the rate of growth of demand is less than the discount rate, then an initial period of accumulation may occur. However, before reaching a steady state, the rate of exploitation of the aquifer will become greater than the rate of recharge. These results do not consider the possibility of an alternative source. Amigues et. al. (1995) do take into account surface water by including hydrologic connections between ground and surface water so that groundwater withdrawals also reduce surface flows. They find that it may be optimal, with a growing population, to let the aquifer replenish because it allows for welfare increasing inter-temporal transfers of surface water flows. We examine the impact of increasing demand when surface water is not connected to the aquifer, but exists as an alternative source of water supply.

Finally, uncertainty should play a role in deciding to conserve groundwater.[8] There may

---

[7] For example, improvements in pumping technology, a permanent reduction in energy prices, or improvements in desalination technologies may permit enhanced utilization of groundwater resources.

[8] This paragraph draws heavily on Conrad and Clark, "Natural Resource Economics", 1987, pp. 209-214.

be uncertainty in recharge or in the costs associated with the alternative source. Both the uncertainty in recharge and in the cost of surface water stem from uncertainty in the amount of precipitation over time. Such uncertainty can be interpreted as being linked with the demand for irrigation, which only comes into play when surface water is scarce. Adding such uncertainty would yield results similar to the deterministic case if demand is linear. Stochastic dynamic solutions methods are similar to certainty equivalents when objective functions are quadratic and the dynamics are linear. The introduction of a concave utility or/and irreversible investment (or disinvestment) may, however, increase the importance of uncertainty. Concave utility introduces risk aversion, which modifies the optimal path by introducing a desire to smooth consumption. Uncertainty in recharge rates and surface water costs is expected to yield a higher steady state equilibrium (conservation). However, the introduction of irreversibility is likely to yield the intuitive result that some groundwater needs to be conserved "just in case". The argument is similar to the one found in the literature on irreversible development. While potential visitors of a park or landmark might not be interested in visiting the place today, they may be uncertain about their future demand and might be willing to pay some amount to keep the *option* of future visits.[9] In such cases, stochastic dynamic programming methods may yield results different from the certainty equivalent approach and uncertainty becomes important in the decision of how much to conserve.[10]

Uncertainty issues are not considered further in this dissertation because we believe that interesting results can be obtained without this complication. Adding uncertainty to the model will be a useful extension for future research.[11] Here, the focus is on surface and

---

[9] This concept was first introduced in 1964 by Burton Weisbrod, who referred to it as "option value".

[10] Henry (1974) shows that development might be found to be optimal using the certainty equivalent approach whereas it might not be optimal using stochastic dynamic programing methods; i.e. it is not sufficient to replace variables with their expected value to find optimal resource use. See also Arrow and Fisher (1974).

[11] We keep in mind, however, that uncertainty may justify limiting the use of water to an amount equal to recharge (MSY) before the aquifer is fully depleted. Also, even though the problem of irreversibility is not dealt with explicitly, we take it into consideration by assuming that the aquifer will never be depleted beyond the point of geological breakdown.

groundwater cost differentials and increasing demand. We examine the cost differentials when there is no lumpy investment and focus on marginal costs. Our analysis with one treatment plant gives enough intuition to address the issue of lumpy investment (for example the possibility of adding a treatment plant in a different location) by looking at the channels through which the assumption of no investment affects our solution. Investment in new surface water treatment plants will greatly modify our results in that it will most likely prevent usage of the full groundwater stock and reduce maximum subsidence.

4. Place in the Economic Literature

Our problem is one of choosing between two sources of the same good when they have different costs. A large number of studies address the issue of resource substitution in the case of *non-renewable* resources (Hotelling,1931; Heal,1976; Kemp and Long,1980; Lewis,1982; Dasgupta and Heal,1979 and 1982). The general result is that, when there are several deposits with different extraction costs, the cheapest deposit should be used first and fully depleted before the next cheapest substitute is utilized.[12] The result also obtains when the alternative is between depleting a non-renewable resource and using a backstop technology (Dasgupta and Heal, 1979).[13] The intuition behind this result is that the money saved from exploiting the low-cost deposit can be invested productively.

It is surprising that the question of which resource should be used first, and according to what criteria, has received so little attention in the *renewable* resource literature. Clarke and Shrestha (1987) suggest that the results obtained for exhaustible resources do not generalize to renewable resources. These authors examine the choice between a renewable resource whose rate of renewal is dependant on the current level of the stock, and a backstop substi-

[12] The result is derived when marginal costs of extraction are constant. Kemp and Long (1980) note that the marginal costs must be constant in terms of utility forgone to extract the resource.

[13] A "backstop technology" refers to a method of producing a product rendering equivalent services (a "backstop substitute"), unlimited in supply, and whose price is exogenously determined. For example, alternative energy sources such as solar energy which can replace fuels without depleting natural deposits.

tute. They find that conjunctive usage of the renewable and the backstop is economically optimal under certain conditions, in particular when the price of the backstop changes over time.

The distinction between *renewable* and *non-renewable* resources is not always clear. Indeed, exploration efforts add to the stock of an exhaustible resource, whereas rates of use permanently greater than rates of recharge exhaust (deplete) a renewable resource. The distinction is especially ambiguous in the case of water resources. Groundwater is usually best described as a renewable resource whose rate of renewal is independent of the amount of water in storage. However, in some cases, the rate of recharge is so small that groundwater is virtually an exhaustible resource. Hence, it is unclear which theory of resource use and substitution applies in the choice between ground and surface water. The geology of the aquifers under consideration will determine whether groundwater is a renewable resource or not.

The decision to use ground or surface water and the possibility of conjunctive usage was first examined by Burt (1964).[14] Like most of the research on water use and groundwater resources, it was directed to explaining agricultural usage. This problem is similar to an inventory problem, as the aquifer can be used as a storage facility. In Burt's analysis, the amount of groundwater to be pumped annually is a random variable dependent upon the stochastic monthly surface supply. He concludes that, where surface water is available, surface water is used first and only supplemented by groundwater if surface water is not sufficient. Ground and surface water are best used conjunctively, but surface water is always used as much as possible and groundwater serves as a buffer to complement surface water in periods of below average rainfall.[15] In this case, problems of groundwater depletion are

---

[14] Most of the research that looks at integrated management of surface and groundwater is found in the engineering and hydrology literature. Buras (1972) studied the best exploitation of a water system consisting of a main surface reservoir and underground storage in the Lake Kinneret area in Israel. The possibility of artificial recharge is explicitly modeled.

[15] On the buffer role of groundwater in agriculture, see also Tsur and Issar (1989), Knapp and Olson (1995).

largely irrelevant as irrigation water is untreated. Consequently, available surface water is cheaper than groundwater, since groundwater is more costly to pump.

In this dissertation, the focus is on the possibility of groundwater depletion where surface water is available. It is relevant for municipal supply, where surface water is not a first choice. For public supply, groundwater is of acceptable quality, located at reasonable pumping depths, is cheaper than surface water because of lower treatment costs.[16]

Many studies have looked at optimal groundwater use in the presence of a pumping externality. The case of salt water encroachment has also been examined. However, the subsidence externality has not been theoretically addressed in economics. Moreover, models of optimal groundwater depletion have not taken into consideration the presence of alternative sources, and very little has been written on optimal use of water resources for municipalities. An article, which touches on the problems associated with the mining of aquifers in cities is by Thomas and Martin (1989). They argue that "growth in willingness to pay for water in cities in metropolitan areas may easily outstrip growth in the direct costs of groundwater extraction in the short term" and that "for an urban area, the decision to proceed with aquifer mining needs to be taken in the context of the array of alternative resources". They suggest the use of an optimal control model of aquifer development in which environmental costs caused by aquifer mining reduce the net benefit to consumers. This environmental cost is the drying of wetlands when water tables fall below a certain level. The model they suggest includes the possibility of an array of alternative sources of water that reduce environmental costs when used instead of groundwater. Thomas and Martin illustrate these issues with the cases of Perth (Australia), Tucson (Arizona), and Bangkok (Thailand). Although subsidence is a phenomenon usually related to the mining of confined aquifers, while the drying

[16] For the same reason, Burt assumes that urban supplies are always satisfied by groundwater reserves (p.65). This assumption allows him to treat urban water use as exogenous in his derivations. Using this result we can treat municipalities as the main user of groundwater and assume that agricultural demand is satisfied entirely by surface water. Note that, in future work, we could incorporate the results concerning the buffer role of groundwater in agriculture by treating irrigation use as causing a stochastic reduction in recharge when below normal rainfall occurs.

of wetlands is related to the dewatering of water table aquifers, both phenomena represent an environmental cost. Subsidence causes environmental damage in the sense that it causes disruption to the natural state of nature through increased flooding and loss of land. We follow their lead using subsidence as the environmental cost, give theoretical structure to the problem, and explore the optimal use of groundwater in meeting urban demand.

Previous studies of water demand have focused on estimations of water demand elasticities, and evaluation of pricing mechanisms (Gibbs,1978; Danielson,1979; Foster and Beattie,1979; Martin and Thomas;1986; Griffin and Chang,1989; Renzetti,1992; and Martin and Wilder,1992). Most studies use linear or log-linear forms to estimate price elasticities, which constrains elasticities to be proportionately decreasing with price (linear form) or constant (log-linear form). Some studies have used flexible forms to allow for variation in elasticity, but these forms are inappropriate to use in dynamic simulations, as they involve a large number of parameters, causing computational difficulties and resulting in poor out of sample forecasts. Our contribution to this literature is to introduce estimates for a Stone-Geary utility function, which has not been used in water demand. We also use estimation techniques more appropriate to panel data.

Policy issues arise when the use of a perceived cheaper source of water needs to be restricted. The presence of external costs causes overextraction if the resource is exploited by more that one entity. Different policies allow costs to be internalized, but how should the groundwater be allocated between different areas? The problem is similar to one of allocating a limited amount of pollution rights, for which the benefits of tradable permits was clearly established in the literature (Stavins, 1998; Crocker, 1966; Dales, 1968; Montgomery, 1972).

5. Outline of the Thesis

The next chapter sets up the long-term water resources management problem faced by

the city planner of a growing city where subsidence occurs with groundwater overdraft, and where surface water is a feasible alternative to groundwater. The spatial allocation problem is solved first. We show that cost minimization implies that if x percent of the city population needs to be serviced with surface water, it should be the x percent closest to the water treatment plant. Results from the spatial allocation problem are then incorporated into the general social planning dynamic optimization problem. The objective is to maximize the present value of net benefits from water for the city, using a representative consumer. Surface water is assumed to be a perfect substitute in consumption to groundwater and, for simplicity, it is assumed that each consumer receives the same amount of water regardless of location.

We formulate the problem using the dynamic programming method of Bellman. The theoretical structure of the model shows that, as demand grows, the water level of the aquifer will go down and subsidence will occur until the social marginal cost of groundwater reaches the marginal cost of surface water. When this occurs, both sources will be used as long as their social marginal costs stay equal. Because of increasing returns to scale in surface water treatment, there will be a period during which water levels will increase and subsidence stops. However, decreasing returns to dispersion will eventually outweigh the economies of scale and as the marginal cost of surface water increases, groundwater levels will again decline. If new treatment plants cannot be built at different locations, subsidence will eventually resume as the groundwater level falls below its historic low. Subsidence will continue until the aquifer reaches its minimum pressure and at that point, groundwater use will remain equal to recharge.

The third chapter introduces a demand framework that can be used to model consumer surplus in the computer simulations. The properties of the Stone-Geary function are well suited for a study of water demand. The function satisfies global concavity properties and

consumer surplus is measured in a simple way. We evaluate the ability of the Stone-Geary function to provide parameters that can be used in the simulations. Choosing functional forms to specify demand in structural dynamic models involves a trade-off between flexibility and parsimony. Flexibility allows elasticities to change over time, while parsimony allows for simplicity, interpretability, and better out-of-sample predictions. The Stone-Geary form uses only two economically meaningful parameters. We show that it also provides flexible price elasticities.

Our empirical study uses panel data on Texas Communities in the early eighties. We estimate a flexible functional form, the Generalized Cobb-Douglas (GCD), using a random effect estimator and compare these results to those based on the Stone-Geary. The Stone-Geary estimates of price elasticities are lower but vary as much as the flexible form results. They are sensitive to seasonality and decrease with price. The Stone-Geary also provides us with an estimate of a water use threshold below which demand is not responsive to price changes. Despite its conditional nature, this parameter may be useful in the design of water conservation policies.

The fourth chapter is devoted to numerical simulation. Information from Houston is used to parameterize the model. All functional forms are specified. The transition equation for water levels is estimated using well data and water use data for Houston. The transition equation for the minimum historic water level, which determines subsidence, is approximated with a differentiable equivalent. The growth of total population is based on census data and published projections. The spatial configuration of the metropolitan area is based on a study of changing urban densities in Houston. The surface water cost function with spatial variables is borrowed from the literature. The pumping cost function is parameterized on the basis of engineering information. Surface water and groundwater cost functions are recalibrated to be consistent with information on cost differentials for Houston. A func-

tional form for the subsidence cost function is assumed so that it matches results of a study for the Houston-Baytown area between 1950 and 1973.

Dynamic programming and numerical techniques are used to solve the Bellman equation derived in chapter 2. Lower and upper bounds for water levels are specified using geological characteristics of Houston's aquifer. To obtain a time-invariant state it is assumed that population reaches a maximum in the year 2050, the beginning of a time invariant state. The value function is numerically approximated with Spline techniques. Backward induction is then used to determine the full paths of the control and state variables.

The fifth chapter deals with applied policy issues using a case study of the Houston-Galveston region. We present the case of Houston and analyze the data using the numerical findings of the previous chapter. Groundwater regulation plans of the Houston-Galveston Subsidence District (HGCSD) are presented and discussed. It is argued that the first plan in 1985, in which different zones are given deadlines to provide most of their water through surface systems, did not take into account potential gains from coordination between adjacent areas, which could reduce the average cost of the alternative surface water provision. The cost of restricting groundwater use varies with the timing of groundwater restrictions and their relationship to the spatial distribution of the regulated zones. We argue that initiatives by some Municipal Utility Districts led the HGCSD to revise its plan in 1992 in a manner more in line with a cost minimizing path. Furthermore, the policy to allocate a fair share of the available groundwater to each area is not optimal because it does not account for differences in the cost of providing water from alternate sources. A system of tradable permits could increase efficiency if there is enough hydrologic connection between the zones.

CHAPTER 2

The Theoretical Model: Methodology and Findings

The spatial allocation of surface water is an important determinant of its cost. Surface water cost differentials among different areas in a city dictate the optimal spatial allocation of surface water. We find this optimal allocation and use the results to specify minimum surface water costs in a social planning model for the entire city, where costs differentials in ground and surface water determine optimal groundwater levels and subsidence. We develop the general structure of benefits, costs, and transition equations, and set up the social planner's optimization problem. We present graphically expected results with and without the option of building treatment facilities in different areas. The no-investment results will be verified in the numerical analysis of chapter 4.

1. Optimal Spatial Allocation of Surface and Groundwater Use

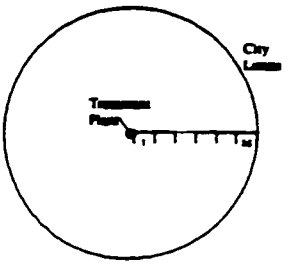
To solve the problem of finding the optimal share of surface water and groundwater to be used in every period for the entire city (or, equivalently the optimal amount of land subsidence), we need to determine how to allocate the use of surface water and groundwater among different areas of a city. The problem of choosing groundwater or surface water depends on the relative cost of these sources. Water supply systems can be separated into three components: (1) acquisition facilities, (2) treatment facilities, and (3) delivery systems (Clark and Stevie, 1981). Delivery systems can be divided into two parts, the conveyance or *transmission* system from the treatment plant to the service area, and the *distribution* system within the service area. The latter is the same for surface and groundwater and therefore does not impact cost differentials. Groundwater is generally decentralized if the aquifer lies

under the whole city and the groundwater needs minimum treatment. Consequently, transmission costs are generally zero for groundwater and the costs of supplying groundwater do not depend on the spatial allocation of groundwater usage. This feature allows us to ignore groundwater costs in looking at the optimal spatial distribution of groundwater and surface water in different areas of the city. In contrast, surface water production and treatment is usually centralized. Raw water is stored in large reservoirs and needs to be transported, while the treatment of surface water involves large fixed costs. Transmission costs for surface water are large and increase with distance. Therefore, the spatial allocation of surface water among different areas in the city becomes an important determinant of costs. The question of which parts of the city should use surface water first needs to be addressed before we can model surface water costs in the general social planning problem.

The social planning problem is greatly simplified if there exists a simple spatial allocation rule for surface and groundwater, and the rule does not determine the value of the optimal control variables in the social planning model. In fact, the spatial problem can be solved first, independently of the optimal control problem, for given aggregate water supplies. The results can then be used to solve the optimal control problem for the whole city without using complicated spatial modelling.

Imagine a circular city with a single treatment plant for surface water at the center.[17] The city is divided into M areas of equal width indexed from $m = 1$ to M according to their distance from the treatment plant, area 1 being the closest and area M the furthest.

[17] The assumption of a circular city is convenient but not necessary. Any form is possible as long as it allows one to divide the city in areas closest or furthest from the treatment plant. Also, results generalize to more realistic cases in which the treatment plant is not centrally located. It is important, however, that the treatment plant be placed close to the most populated area. Although surface water plants are not generally located close to populated area, the infrastructure put in place at the time of the building of the plant is designed to service the most populated area as close as possible to the plant. For example, most cities were developed from a surface water source and the infrastructure was first developed to serve the center. As population spreads, it spreads away from this original infrastructure. Considering this, and the fact that most cities have spread far beyond their original city limits, we believe that this assumption is not too unrealistic and will allow for general results.



The city has both groundwater and surface water and the consumers are indifferent between the two sources.[18] Each area m consumes s_m of surface water and g_m of groundwater. Each area has a population N_m, with

$$\sum_{m=1}^{M} N_m = N.$$

We assume that per capita consumption is the same in each area.

$s_m + g_m = N_m\overline{w}$ for all m.

We do not assume any specific population distribution except that $N_{m-1} \geq N_m$ for all m. This allows us to ignore cases when diseconomies of dispersion will never offset economies of scale. Let the average cost of providing surface water be $f(m)$, an increasing function of m.[19]

Let x be the proportion of surface water to total water use that the city wishes to produce. The city planner needs to find the best allocation of surface water among the different areas so that total water costs are minimized. Since only surface water transportation costs are affected by the spatial allocation of water, the planners problem becomes one of minimizing surface water transportation costs. Given that a proportion x of the water is surface water, the planner's problem is to minimize surface water transportation costs

$$C = \sum_{m=1}^{M} s_m f(m)$$

subject to, for all m, $\left(\sum_{m=1}^{M} s_m\right)(1-x) = x\sum_{m=1}^{M} g_m$, $s_m + g_m = N_m\overline{w}$, and $s_m \geq$

[18] This allows us to ignore the benefit side in the optimal spatial allocation problem and concentrate on cost minimization.

[19] f is not a function of the quantity of surface water since we ignore the impact of returns to scale. This has no bearing on the solution as long as population density does not increase with distance.

$0, g_m \geq 0$. Where the first two constraints can be combined to get $x\left(N\overline{w} - \sum_{m=1}^{M} s_m\right) = (1-x)\sum_{m=1}^{M} s_m$.

Proposition 1 : *If for some $m \in (1, M)$, $s_m > 0$, then $g_{m-1} = 0$; i.e., if area m uses some surface water then it must be true that area $m-1$ relies solely on surface water and uses no groundwater*

Proof of proposition 1:

The Lagrangian for the cost minimization problem is

$$L = -\sum_{m=1}^{M} s_m f(m) + \lambda\left[x\left(N\overline{w} - \sum_{m=1}^{M} s_m\right) - (1-x)\sum_{m=1}^{M} s_m\right]$$
$$+ \sum_{m=1}^{M} \mu_m s_m + \sum_{m=1}^{M} \varphi_m\left(N_m\overline{w} - s_m\right)$$
$$= -\sum_{m=1}^{M} s_m f(m) + \lambda\left(N\overline{w}x - \sum_{m=1}^{M} s_m\right) + \sum_{m=1}^{M} \mu_m s_m + \sum_{m=1}^{M} \varphi_m\left(N_m\overline{w} - s_m\right),$$

where λ, μ_m,and φ_m all m are Lagrange multipliers.

The first order conditions and complementary slackness conditions for $m = 1, ..., M$ are:

$-f(m) - \lambda + \mu_m - \varphi_m = 0,$

$\mu_m s_m = 0, \mu_m \geq 0, s_m \geq 0,$

$\varphi_m\left(N_m\overline{w} - s_m\right) = 0, \varphi_m \geq 0, N_m\overline{w} \geq s_m.$

Then, for any $m_1 < m_2$, we have $\mu_{m_1} - \varphi_{m_1} - f(m_1) = \mu_{m_2} - \varphi_{m_2} - f(m_2)$, which implies

$f(m_2) - f(m_1) = \mu_{m_2} + \varphi_{m_1} - (\varphi_{m_2} + \mu_{m_1}) > 0$, since $f'(m) > 0$.

Hence,

$\mu_{m_2} + \varphi_{m_1} > \varphi_{m_2} + \mu_{m_1} \geq 0$

$\Rightarrow \mu_{m_2} > 0$ or $\varphi_{m_1} > 0$ or both

$\Rightarrow s_{m_2} = 0$ or $s_{m_1} = N_m\overline{w}$, i.e. $g_{m_1} = 0$ or both.

We conclude that either $g_m = 0$ for all m or $g_m = 0$ for a range of m up to some $\overline{m}$. At $\overline{m}$ we could have $\mu_{\overline{m}} = \varphi_{\overline{m}} = 0$ but beyond $\overline{m}$ we must have $s = 0$ and $g \geq 0$.

Proposition 2 : *If x is the proportion of surface water to total water use in the city, then the distance to cover for the provision of surface water equals the distance from the center necessary to reach a population of $x.N$, where N is the total population in the city.*

Proof of proposition 2: This follows directly from proposition 1 and the assumption of uniform per capita consumption. If $x > 0$, then at least one area m^* uses surface water. Let m^* be the area furthest from the center that uses surface water. All areas indexed $m < m^*$ must be using exclusively surface water by proposition 1. Since we assumed that per capita consumption of water is uniform, the proportion of individuals using surface water must be equal to the proportion of surface to groundwater use, so we must have $\sum_{m=1}^{m^*} N_m = x.N$. The distance from the center to the outer edge of m^* is the distance necessary to reach a population of $x.N$. Let us denote this distance by d_x. It is clear that any smaller distance would reach less than $x.N$ and any longer distance would reach more than $x.N$.

This result does not depend on the population distribution among the different areas, except for the assumption of non-increasing density. However, the actual distance d_x (distance necessary to reach a population of $x.N$) does depend on population distribution. It is necessary to find a functional form that relates d_x to the population characteristics of the city. The functional form must have the following characteristics: for any population distribution, it must always be true that $d_0 = 0$ and $d_1 = dM$, where d_0 is the distance necessary when no surface water is used ($x = 0$), d_1 is the distance necessary when only surface water is used ($x = 1$) , and dM is the distance from the center to the city limits. If density is uniform, then $d_x = x.(dM)$. For example, to supply one third of water from surface water, one needs to go one third of the total distance from the center to the city limits. If the population is concentrated at the center (i.e., if population density drops immediately to zero) then the distance d_x is zero and the spacial allocation becomes irrelevant. A density gradient λ is often used to describe how population density drops from the center when density is assumed to decrease exponentially. Several studies have shown that an exponentially decreasing den-

sity can approximate population distribution in many cities (Mills, 1972; Edmonton, 1975). λ can be interpreted as the constant percentage drop in density from the center. A high λ indicates a very centralized city and a low λ indicates that the city is spread out.[20] Using such a density function, the uniform density corresponds to $\lambda = 0$ and the perfectly centralized city to $\lambda = \infty$. The assumption of exponentially decreasing density will be incorporated in our simulation of chapter 4. We will relax the assumption of circularity and let the geometry of the city be some portion of a circle (angle of integration will be $\overline{\Theta}$).

We can easily incorporate our function d_x into the social planning model to calculate the cost of providing surface water to the city. This allows us to eliminate any reference to different areas of the city in the social planning model. The transportation cost of surface water will only depend on the ratio of surface water to total water use, and the density gradient.

2. The Mathematical Model: Dynamic Programming Approach to the Optimal Control Social Planning Problem.

The social planning problem is one to optimally choose the level of groundwater use and surface water use in order to maximize net benefits from water consumption. We set up a deterministic optimal control model of groundwater and surface water choice where there is both a pumping and a subsidence externality and where demand increases over time. We choose a discrete time dynamic programming approach to the problem, a technique that is well suited to numerical approximation. The general setup is similar to traditional groundwater management models where pumping costs depend on the water level in the aquifer, which in turn depends on the amount of water withdrawn in previous periods and

[20] The validity of the gradient approach has been questioned because very different results are obtained when empty land is taken into account. The net density gradient, which is corrected for vacant land, tends to measure how population density drops along some major highways (Mieszkowski and Smith, 1991; Mieszkowski and Mills, 1993). This distinction does no affect our results since we are not interested in the density per say but how far one needs to go to reach a certain proportion of the population.

the rate of recharge. If the rate of extraction is greater than the rate of recharge, water levels decrease and it becomes more expensive to pump groundwater. We add surface water as an exogenous alternative supply source. The quantity of surface water is unconstrained and renewed every year, independently of the state of the system. We consider only the cost of production, treatment, and transmission of surface water, assuming that surface water is not purchased. Although rights to surface water exist in Texas and raw water can be traded, we abstract from this complication, noting that the addition of a price of surface water to our model may be a useful direction for future research.[21]

Although the cost of providing surface water varies spatially, we assume that there is a single price for water in the city, so consumption patterns are not affected spatially, through price differences. In a social planning problem, price does not enter explicitly, but marginal costs and marginal benefits together determine the optimal use of water. However, constraining the social planner to allocate the same amount of water to all residents and implicitly charging each consumer an identical price has important implications. It means that the social planner cannot vary the amount of water across consumers, even if this would increase total net benefits. Uniform water prices are commonly observed in municipalities where price is based on the average cost of providing water to the entire city. The assumption of a uniform price is not unrealistic and can be interpreted as an equity constraint.

2.1 Model Variables

2.1.1. State Variables

There are three state variables, two of them are endogenous, and population determined exogenously.[22] The state of the system is completely described by the first two interdepen-

---

[21] An exogenous price for surface water could be incorporated as in Clarke and Shrestha (1987). The price of raw surface water is similar to a nonautonomous price of a backstop substitute. This becomes interesting when we recognize that planners are likely to be uncertain about future prices. Since we are leaving uncertainty aside for now, we do not lose much by ignoring the price of raw surface water.

[22] Since we assume that surface water is renewed every year independently of past decisions, we do not need a state variable for surface water.

dent state variables:

H_t is the hydraulic head at time t. Since subsidence mostly occurs with confined aquifers, the head is the piezometric surface of the aquifer, i.e. the average water level in wells.[23] H_t is constrained to be within the following bounds:

$$\underline{H} \leq H_t \leq \widetilde{H}$$

where $\widetilde{H}$ is the maximum head, when the aquifer has reached maximum capacity, and $\underline{H}$ is the minimum head, determined by geological characteristics of the aquifer. If usage is less than the recharge and $H_t = \widetilde{H}$, the difference is lost as runoff. The lower the value of H, the higher the pumping costs.

M_t is the minimum historic hydraulic head level in the aquifer. It determines the cumulative amount of subsidence. We start in a state when $H_0 = M_0$. H and M coincide as long as H does not increase. M is constant when H is greater than M. The boundary conditions for M are:

$$\underline{H} \leq M_t \leq H_0$$

N_t is the total population to be served in the urban area at time t. Contrary to the other two state variables, it is not driven by the values of the controls. Following Amigues et. al. (1995), growing demand is incorporated through population growth. The population level is a state variable that varies exogenously over time. The evolution of N_t drives the model by shifting demand over time.[24]

2.1.2. Control Variables:

Two control variables determine the evolution of the system. The planner chooses g_t the quantity of groundwater to be withdrawn each period t from the aquifer and s_t, the quantity

[23] In a confined aquifer, water is under pressures greater than atmospheric pressure. The water level in the well reflects the pressure in the aquifer. The greater the pressure, the higher the water in the well. This is different from unconfined (or "water table") aquifers where the pressure is the same as the atmospheric pressure at the surface, and therefore where water levels in the wells are also the water level in the aquifer.

[24] We could imagine that the index t be replaced by a population index. Although we leave t as the time index because we need a normal dynamic structure, and include N as an exogenous state, it is, in effect, not time but levels of populations at different times that drives the system.

of surface water to be used each period. As noted above, non-negativity constraints apply to both variables, so:

$g_t \geq 0$ and $s_t \geq 0$ for all t.

2.1.3. Other Variables:

The city is completely characterized spatially by values of $\overline{\Theta}$, a parameter that describes the geometry of the city, the density gradient (λ), the population density at the center (K), and the distance from the city center to the city limits (l). There exists an exact relationship between these variables and N_t. Implicitly, λ, l, K,and $\overline{\Theta}$ are functions of N_t and some of these variables will need to change over time as N changes. In the simulations, we allow l and λ, and K to vary with time (population), and keep $\overline{\Theta}$ constant. All other parameters are constant. In particular, the discount factor β is assumed to be constant. Sensitivity analysis will be carried out to evaluate the importance of the magnitude of some of the autonomous parameters.

2.2 Model Functional Equations

2.2.1. Transition Equation for the State Variables:

The transition equation for groundwater levels follows the specification used by Brown and Deacon (1972) and Gisser and Sanchez (1980).[25] The first state equation specifies the evolution of H. All withdrawals are consumed so there are no return flows into the system. Water levels are linearly related to groundwater use[26], with water level in the aquifer following the following difference equation:

$$H_{t+1} = H_t + A - Bg_t,$$

---

[25] Brown and Deacon (1972) use feet of lift instead of water levels to describe the state of the system. These two approaches are equivalent. Gisser and Sanchez (1980) also incorporate a return flow coefficient in the analysis. We assume no return flows.

[26] The relationship between water withdrawal and head has been shown to be non linear. The instantaneous effect of water withdrawal become less and less as more is withdrawn. However, this relationship concerns mostly the instantaneous head change due to the cone of depression caused by pumping. When it comes to permanent changes in water levels, the relationship is approximately linear.

where A is the annual natural recharge in feet of head and B is a factor of dimensionality in feet per acre foot, so A/B is annual natural recharge in acre feet.[27]

$H_t \leq \overline{H}$ for all t, so that if $H_t + A - Bg_t \geq \overline{H}$, then $H_{t+1} = \overline{H}$. Therefore, the transition equation holds as long as $g_t \geq \frac{A+H_t-\overline{H}}{B}$, otherwise the head level stays at $\overline{H}$.

The second state transition equation specifies the evolution of M. There is no change in M unless $M_t = H_t$ and groundwater levels are going down. In this case the two state variables reduce to one. In mathematical terms we have:

$M_{t+1} = M_t$ when $H_t + A - Bg_t \geq M_t$ or, equivalently, if $g_t \leq \frac{A+H_t-M_t}{B}$;

$M_{t+1} = H_{t+1} = H_t + A - Bg_t$ otherwise.

The transition equation for population N does not include any control variables. The state of the population influences the system exogenously. Population increases until it reaches a maximum of $\overline{N}$ at a given future date, which will be the starting point of the time-invariant system. Although the endogenous state variables H and M can continue to evolve after that date, the system is stationary as a function of $H$and M. We assume an arithmetic population growth function:

$N_t = N_0 (1 + nt)$

2.2.2. Gross Benefit Function:

Gross Marshallian consumer surplus is used as the measure of benefits. As the fraction of total income spent on water is small, income effects are small, so a Marshallian surplus measure can be used as an approximation of welfare benefits.[28] $W(w_t)$ is per capita gross consumer surplus derived from a per capita annual water consumption of $w_t = \frac{s_t+g_t}{N_t}$. Benefits depend solely on the level of water consumption, as surface and groundwater are

---

[27] This formulation follows Brown and Deacon (1972). For confined aquifers, B can be approximated using $\frac{\ln(r_0)}{2\pi Kb}$, where r_0 is the specific radius of influence, K is hydraulic conductivity, b is aquifer thickness (Kb = transmissivity).

[28] Consumer surplus is commonly used to measure benefits in the literature [see for example Dasgupta and Stiglitz (1981), Clarke and Shrestha (1987), Amigues et al. (1995)]. For a discussion of the appropriateness of Marshallian surplus measures, see Willig (1976), Hausman (1981), Blackorby and Donaldson (1985).

perfect substitutes.[29] Following Amigues et. al. (1995), we assume that all individuals are identical so total consumer surplus can be calculated on the basis of per capita consumption multiplied by the population. Preferences exhibit diminishing marginal utility in water consumption so $W_w > 0$ and $W_{ww} < 0$.

2.2.3. Cost Functions : Groundwater

Groundwater costs include pumping cost and the environmental cost of subsidence. Since the local distribution system for groundwater is also used for the distribution of surface water, local distribution costs can be ignored in the analysis since they do not affect cost differentials.

Pumping costs depend on the height from which water must be pumped (the lift), which is the difference between land surface (S_L) and water levels in the well (H). Although there is subsidence in the model, the change in land surface level is not large enough to affect pumping costs. The problem is simplified by assuming a constant elevation land surface in the pumping cost function. We use $C_p(g, H)$ to describe the cost of pumping g units of groundwater when the head level is H, for a given land surface elevation. Pumping costs consist mostly of energy costs. These costs depend on the price of energy and the amount of horsepower needed to lift water. Horsepower requirements are directly proportional to the size of vertical lift and inversely proportional to the pump efficiency. We assume that the rate of pumping does not vary, so to withdraw a greater quantity of water, pumps work longer. The function can be written as $C_p(g, H) = g[c_0 - c_1 H]$, a formulation which is common in the literature.[30]

---

[29] Municipal water users usually do not care whether water comes from the ground or from surface sources. Although, examples can be found where surface water projects were abandoned because the public does not recognize surface water as being a perfect substitute to groundwater. For example, the Applewhite reservoir project for the City of San Antonio was abandoned in the early nineties partly because of the resistance of residents to using treated surface water instead of the water from the Edwards aquifer. We abstract from such considerations, mostly for reasons of simplicity and because such behavior has not been noticed in Houston, which is our main case study.

[30] Martin and Archer (1971) and Gisser and Sanchez (1980) use a similar assumption. Brown and Deacon (1972) assume that pumping costs are increasing in lift. This occurs if new pumping equipment is needed or wells need to be re-drilled when depth increases. In this model we simplify the analysis to look at cases when such investments are not necessary.

We consider only permanent costs of subsidence. Although the costs of structural damage can be large, it is small compared to the accumulated cost of lost land and increased flooding. We ignore the costs related to the rate at which land subsidence occurs and focus on costs associated with the cumulative amount of subsidence. We also ignore costs due to a permanent loss in storage capacity in the aquifer.[31]

We use $C_s(M)$ to describe the cost of subsidence when the minimum historic head is M. Subsidence is largely irreversible. When water levels increase, subsidence is stopped (but not reversed). Only if the water level falls below its lowest historical level will subsidence resume.

Data on the cost of subsidence is limited and we found no published functional form that relates subsidence costs to groundwater withdrawals. However, a commonly accepted first approximation is that the amount of subsidence is linearly related to the drop in head levels below the minimum historic head. A study by Warren et al. (1975), gives some indication of the magnitude of subsidence costs and how they have increased in time as water levels dropped. The evidence points toward a subsidence cost function with increasing marginal cost $C_s' > 0, C_s'' > 0$.

2.2.4. Cost functions: Surface Water

The cost structure of surface water is independent of groundwater levels. Surface water costs include acquisition costs such as pumping and storage, treatment costs, and costs of transporting the water to its point of use (called transmission or conveyance costs). Transmission costs constitute a large part of surface water costs and depend on the spatial allocation of the water. It has been shown that total surface water costs exhibit economies of scale when the quantity of surface water per unit area increases, but diseconomies when service area alone increases (Ford and Warford,1969). We exploit this feature and use a

[31] Since the loss of storage capacity is small compared to total capacity in the aquifer, such costs are negligeable.

functional form that incorporates these two effects in formulating total acquisition, treatment, and transmission costs. The functional form we use is from Clark and Stevie (1981), where there are economies of scale in treatment and transmission, but diseconomies in dispersion.[32] Acquisition and treatment costs exhibit increasing returns to scale until the maximum amount of capacity is added to a single location treatment facility. We force the average cost of surface water treatment to increase sharply after this maximum single location capacity is reached. Transmission costs always increase with distance. We assume that surface water usage is allocated optimally within the metropolitan area, as derived in section 2.1.[33]

$f(s, d_x(.), \overline{\Theta}, K, N)$ is the total cost of surface water treatment and transmission costs when s_t is the rate of surface water use and $d_x(x, \lambda, l)$ is maximum distance that surface water needs to be transported to reach a proportion $x = s/(g+s)$ of the urban population, given a density gradient of λ and a maximum distance to the city limits of l. d_x increases with x and l and decreases with λ.

2.3 The Dynamic Optimization Problem

2.3.1. Objective Function

The social planner chooses g_t and s_t for each period t to maximize the present value of discounted net benefits from water consumption subject to the physical balance equations that determine current and minimum historic water levels. Net benefits consist of the consumer surplus derived from consuming water minus the costs of production. β is the constant discount factor. We assume that groundwater supply is developed, i.e. that all the infrastructure to provide groundwater is in place. All variables are deterministic, and all

---

[32] Recall that distribution costs from the main pipes to the consumer are the same for surface water and groundwater so we ignore them. Transmission costs are different from these distribution costs in that they involve the main pipes from the treatment plant to the smaller distribution system.

[33] Since we assume one location only for treatment, optimal placement of the treatment plant (or group of plants) is also assumed.

costs are variable.

The social planning structural problem (SP) can be stated as follows:

$$\max_{g_t, s_t} \sum_{t=0}^{\infty} \beta^t \left\{ N_t W(w_t) - C_p(g_t, H_t) - C_s(M_t) - f(s_t, d_x(x_t, \lambda), K, l, \overline{\Theta}, N_t) \right\} \quad (2.1)$$

where, as defined above, $w = \frac{g+s}{N}$, and $x = \frac{s}{s+g}$.

subject to the transition equations:

$$H_{t+1} = H_t + A - Bg_t \quad (2.2)$$

$$M_{t+1} = M_t \text{ if } H_{t+1} > M_t \quad (2.3)$$

$$= H_{t+1} \text{ otherwise;}$$

$$N_t = N_0(1 + nt) \quad (2.4a)$$

the non-negativity constraints: $g_t \geq 0$ and $s_t \geq 0$ for all t;

and boundary conditions: $\underline{H} \leq H_t \leq \tilde{H}$; $\underline{H} \leq M_t \leq M_0 = H_0$; $N_0 \leq N_t \leq \overline{N}$.

Note that the boundary constraint on H together with the transition equation for H imply that g is constrained by aquifer capacity so that at all times

$$Bg_t - (H + A - \underline{H}) \leq 0 \quad (2.5)$$

Also surface water is limited to the capacity in place if we do not allow for investments in new treatment plants.

2.3.2. Dynamic Programming Formulation

To solve the dynamic programming problem, we need to find laws of motion that link the state variables in two adjacent periods. Using Bellman' s principle of optimality, at any

period t, the optimal choice of controls for periods $t+1$ and beyond does not depend on states of the system prior to period t, but only on the resulting state at the beginning of period t. We define the value function $V_t = V(H_t, M_t, N_t)$ as the present discounted value of net benefits. V_t is a function of the current state of the system; i.e., current water levels, current subsidence level (which depends on M_t, the minimum historic water levels), and current population. The following functional equation (FE) summarize the dynamic problem:

$$V_t = \max_{g_t, s_t} \left\{ N_t W\left(\frac{s_t + g_t}{N_t}\right) - C_p(g_t, H_t) - C_s(M_t) - f(s_t, g_t, N_t) + \beta V_{t+1} \right\} \quad (2.6)$$

where V_{t+1} is a function of the next period states, given by our transition equations 2.2, 2.3, and 2.4a above. This is the Bellman equation that will be used in the numerical analysis of chapter 4. We will also verify in chapter 4 that our functional forms, constraints, and feasible control sets satisfy the conditions under which the SE and FE problems are equivalent. We now try to get some insights into the analytical results that the model produces by looking at the structure of the first order conditions.

The value of s does not directly affect the state of the system. Therefore, the first order conditions for s do not involve the state variables. To find s_t^*, the optimal amount of surface water at time t given g_t, the first order necessary condition requires that $\frac{\partial V_t}{\partial s_t} = 0$ for any time period t. This implies that

$$\frac{\partial W}{\partial (s+g)/N} = \frac{\partial f}{\partial s} \quad (2.7)$$

s^* that solves the F.O.C. is therefore the amount of surface water for which the marginal benefit from water use (given g and N) equals the marginal cost of providing surface water. This must hold for each period t. The second order sufficient condition requires that

$$\frac{\partial^2 W}{\partial s^2} - \frac{\partial^2 f}{\partial s^2} < 0 \tag{2.8}$$

By assumption of diminishing marginal utility, the first term is negative. A positive second derivative for the cost function would suffice to ensure a maximum. However, the surface water cost function may be increasing at a decreasing rate over a range of values when the decreasing returns to distance do not outweigh the increasing returns to scale. When increasing returns dominate, we need the rate of decrease in marginal cost to be smaller than the rate of decrease in marginal benefit. If this were not the case, the solution to the first order condition would be a minimum, not a maximum. The optimal s would be a corner solution where it becomes optimal to expand surface water until consumers are satiated. This is unlikely to be the case in practice. Henceforth, we assume that the solution for optimal s is always interior.

The first order condition for a maximum with respect to g involves the state variables and therefore is dynamic in nature. In order to get a closed form for the first order condition, we rewrite the transition equation for M as

$$M_{t+1} = M_t + min\left(H_t + A - Bg_t - M_t, 0\right) \tag{2.9}$$

The transition equations (2.2) and (2.9) can be substituted into the Bellman equation (2.6) to get the first order conditions for a maximum at time t.

$$\frac{\partial W}{\partial (s+g)/N} = \frac{\partial C_p}{\partial g_t} + \frac{\partial f}{\partial d_x}\frac{\partial d_x}{\partial g_t} + \beta\left(-B\frac{\partial V_{t+1}}{\partial H_{t+1}} + \frac{\partial V_{t+1}}{\partial M_{t+1}}\frac{\partial M_{t+1}}{\partial g_t}\right) \tag{2.10}$$

We will use numerical techniques to approximate the unknown form of V and its derivatives in chapter 4. For the sake of interpretation, we just need to understand that $\beta B\frac{\partial V_{t+1}}{\partial H_{t+1}}$ is the discounted effect that the current level of groundwater withdrawal has on future net

benefits through its impact on the water level in the aquifer, and $\beta\left(\frac{\partial V_{t+1}}{\partial M_{t+1}}\frac{\partial M_{t+1}}{\partial g_t}\right)$ is the discounted effect that the current level of groundwater withdrawal has on future net benefits through its impact on subsidence, which is equal to $\beta\left(-B\frac{\partial V_{t+1}}{\partial M_{t+1}}\right)$ when $H_{t-1} < M_t$ (subsidence occurs), zero otherwise. When $H_{t+1} < M_t$ (subsidence occurs), then the marginal user cost of groundwater includes subsidence costs. In this case the discounted user cost of groundwater is $\beta\left[-B\left(\frac{\partial V_{t+1}}{\partial H_{t+1}} + \frac{\partial V_{t+1}}{\partial M_{t+1}}\right)\right]$.

The usual intuition applies to this result. The marginal benefit from water consumption must be equal to the marginal cost associated with current groundwater withdrawals, but this time the marginal cost includes the discounted sum of all costs imposed on the future by current groundwater withdrawals through the withdrawal's impact on the state of the system (head level and subsidence). The second order sufficient condition is satisfied with non decreasing marginal costs, which are assumed in the case of groundwater. We will also need to verify that the unknown value function is concave.

Since g and s are perfect substitutes in consumption, and assuming that the constraint on g is not binding, the two first order conditions imply that both sources will be used conjunctively only if their marginal costs, including user costs, are equal. In the absence of this equality, the source with lowest marginal cost is used exclusively.

3. Expected Results: Behavior of State Variables and Controls

3.0.3. Base Model: No Further Investment in Surface Water Treatment

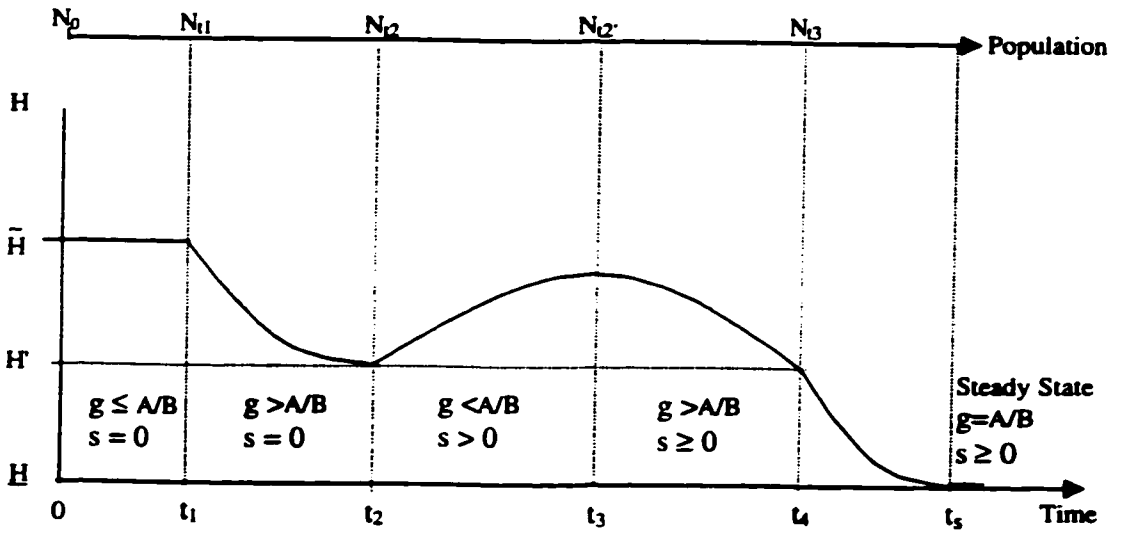
In a steady state, we must have groundwater use equal to recharge, $g = A/B$. The marginal cost of groundwater (MCg) then remains constant and at most equal to the marginal cost of surface water (MCs). If MCg is less than MCs then it must be true that the constraint on groundwater use is binding (the "bottom" of the aquifer is reached). MCg cannot be less than MCs in a steady state, otherwise, less groundwater would be used and water

levels would go up. Therefore, a steady state can only be reached if the marginal cost of surface water is nondecreasing in s. This last point is very important and makes the problem quite interesting. If increasing returns to scale outweigh decreasing returns to dispersion, we should have MCs decreasing as soon as surface water starts to be used. However, as population increases, there must be a point at which decreasing returns to dispersion take over and MCs starts increasing. If capacity cannot be increased, then MCs keeps increasing. Conjunctive use implies that the amounts of groundwater and surface water used are such that $MCg = MCs$, with g below its upper bound. We expect the use of groundwater to decrease as MCs decreases (so that MCg also decreases), and to increases as MCs increases (so MCg also increases). As MCs increases, the use of groundwater will stay above recharge until the water level in the aquifer reaches its minimum level, determined by geological characteristics. How fast that level is reached depends on the relative rates of increases in marginal costs. At this point a steady state is reached, maximum subsidence has taken place, and groundwater use is limited to recharge; i.e., in the steady state, $H = M = \underline{H}$. The rate of surface water use will then increase as determined by population growth and the utility parameters, but groundwater use cannot increase beyond recharge. This scenario is mainly driven by population growth and the surface water cost functions. Graphically, we expect the behavior of the state and groundwater use to be as represented in Figure 2:

$0 - t_1$: The aquifer is being used but has never been depleted. Population is still small enough to be serviced by the amount of groundwater recharge or less, so the head level stays at its maximum and no subsidence has occurred. In this interval, $MCs > MCg = \overline{y}$, where $\overline{y}$ is the minimum groundwater cost; $g \leq A/B$ and $s = 0$.

$t_1 - t_2$: Population, and therefore the demand for water, is growing so that recharge is not sufficient. However, as long as $MCs > MCg$, it is optimal to use more groundwater

Figure 2: Behavior of the System with Single Location Treatment Plant



and start depleting the aquifer; $g > A/B$; H decreases and so does M and, with M the level of subsidence; MCg increases with increased pumping costs and subsidence costs.

At t_2 : $MCs = MCg$, surface water becomes an economical alternative, conjunctive use starts.

$t_2 - t_3$: Economies of scale in surface water dominate diseconomies of dispersion, so MCs decreases as s increases. It is now more economical to use surface water so the use of groundwater decreases. As g decreases, the aquifer is being replenished, marginal pumping costs decrease and subsidence ceases, so MCg decreases too; $g \leq A/B$ and decreasing; s increases; MCs decreases at a decreasing rate because of diseconomies of dispersion, so the rate of reduction of groundwater use and therefore the rate of replenishment must be decreasing too. At t''', MCs is minimized. M stays at the H' level.

$t_3 - t_4$: Increasing cost of surface water delivery takes over, so MCs increases. It is once again worthwhile to use more groundwater and to deplete the aquifer. However, subsidence does not occur until $H = H'$. Until then, the increase in MCg is only due to pumping costs.

At t'''' the level of the aquifer reaches $H' = M$, its historic low.

$t_4 - t_s$: As H decreases beyond H', subsidence costs become once more a part of MCg as M starts decreasing again with H, causing the rate of change of MCg to increase. MCg increases at the same rate as between t_1 and t_2 as more groundwater is used. The rate of increase in groundwater use is such that the increase in MCg equals the increase in MCs.

t_s to ∞ : The aquifer cannot be depleted further due to its geological characteristics. Only the recharge can be used. A steady state is therefore reached for the use of groundwater and the water level in the aquifer. Subsidence is halted, but not before maximum damage has occurred..

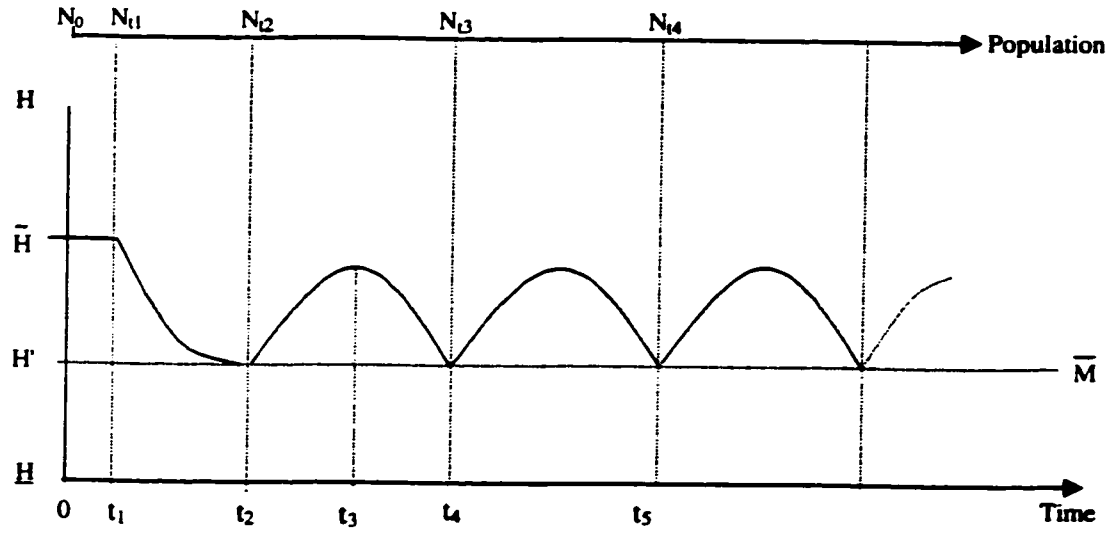
3.0.4. Extension to the Case of Lumpy Investment in Surface Water Treatment.

The preceding results rest on the assumption that there is only one treatment plant, built at the optimal place and ready to operate at t_2. Why would a treatment plant be built at this time? It must be optimal to do so to avoid further increases in subsidence and groundwater pumping costs. Figure 3 describes the anticipated behavior of our state variables when treatment plants can be built in different locations.

The future cost-savings from using less groundwater from t_2 to t_5 must be at least equal to the cost of setting up a treatment plant. If treatment plants could be built every period in a non-lumpy fashion, and if t_2 was indeed the optimal time to start using surface water, then it is obvious that capacity would be increased just enough every period so as to maintain the level of the aquifer at H'. A steady state would be reached at time t_2 with H' as the steady state head. This is a higher steady state than in the no-investment case, meaning that the capacity of the aquifer would not be fully used before the steady state is reached (conservation).

What would be the optimal time to build a second treatment plant? Let us assume constant technology, so treatment plants have the same characteristics and their cost remain

Figure 3: Behavior of the system with Lumpy Investment



constant. If it is optimal to have a treatment plant ready to operate at time t_2, then it must be optimal, for the same reasons, to make a new investment of the same size at time t_4, t_5, etc.. Indeed, at time t_4 and t_5, groundwater pumping costs and subsidence costs would start increasing again and in the same fashion as they would have if the treatment plant had not been built at time t_2. It is worth it to build a new similar treatment plant to avoid the same types of costs as were forecasted at time t_2. The optimal time to build the new plants depends on the growth of demand. An arithmetic growth rate would cause treatment plants to be built at constant intervals, while an exponential growth rate would cause the intervals between construction to become shorter.[34]

[34] As in Manne (1967), we assume that plants come only in one size. If the size of the plant is also a choice variable, and there are economies of scale in building new plants, the problem is more complicated. Hartley and Kyle (1989) analyze this case.

CHAPTER 3

Choosing Demand Parameters: Evaluation of the Stone-Geary Form Using Panel Data Estimation.

1. INTRODUCTION

The estimation of water demand involves a trade-off between parsimony and global regularity of the preference structure on the one hand, and flexibility on the other.[35] Most of the earlier studies in water demand have opted for simplicity, using linear and/or log-linear specifications.[36] The resulting restrictions on elasticities are not problematic as long as price variation is small, and predictions are made within a similar price range. However, with a long-run assumption of increasing water scarcity, linear or constant-elasticity forms are problematic. In running dynamic simulations, it is common to ignore this problem, decide on a simple consumer surplus from a linear demand function, and use elasticities from the empirical literature to parameterize benefits. We believe that it is better to estimate demand and use the same functional form and estimated parameters to model the benefit side of our objective function. Because there is evidence that price elasticity is neither constant nor linearly on the price level, we need a functional form that allows for nonconstant elasticities.[37] Nonconstant elasticity results have been obtained using flexible functional forms, but they are not well suited for specification of benefits in dynamic programing simulations

---

[35] Flexibility usually refers to the ability of a functional form to approximate a twice continuously differentiable function to the second order. The number of free parameters required for this purpose is large. In particular, if the number of exogenous variables is N, a flexible consumer demand function requires $(N^2 + 3N - 4)/2$ free parameters (Diewert and Wales,1988).

[36] Danielson (1979) summarizes studies done before. Later studies include Cochran and Cotton (1985).

[37] Foster and Beattie (1979) obtain non-constant price elasticities using a price exponential model. Griffin and Chang (1990 and 1991) compare linear, log linear, Generalized Cobb-Douglas, Translog, and Fourier functional forms. Only the flexible forms give nonconstant elasticities.

for reasons explained below.

Although flexibility often provides good sample fits, out of sample forecasts suffer if variables trend away from their within sample values. Global curvature restrictions can be imposed on these flexible forms, but, unless flexibility is greatly restricted, the number of parameters needed is still large (Diewert and Wales, 1987). Some methods address the concavity problem by only imposing restrictions locally "over a set of prices where inference will be drawn" (Terrell, 1996), but this does not solve the out-of-sample problem.

Hausman and Newey (1995) used non-parametric methods to estimate exact consumer surplus. Dynamic optimization models usually require consumer surplus, but the consumer surplus must be expressed as a function of control variables. A non-parametric estimation with numerical approximation of a demand function cannot be used for this purpose. The main problem remains one of parameter parsimony. None of these functions can be used in dynamic simulation models, which require the full form of the demand function. In addition, in most flexible form estimation, parameter estimates are functional approximations and often do not have any straightforward economic interpretation. To get better insights on results of dynamic optimization models and their sensitivity to parameterization, it is important to be able to interpret the demand parameters with reference to an underlying structural model.

The present chapter proposes and estimates a functional form that is simple, interpretable, and generates nonconstant price elasticities. We propose the use of a Stone-Geary demand function and compare its performance to a flexible functional form, previously used in water demand estimation, the Generalized Cobb Douglas (GCD).[38] The comparison is based on the estimation of water demand using five years of monthly data from a large number of Texas communities. We find that the elasticities obtained for the Stone-Geary have sea-

[38] In their comparison of several demand specifications, Griffin and Chang (1991) find that the Generalized Cobb-Douglas form (GCD) provides flexibility in the parameters while preserving a relatively nicely behaved demand function.

sonality patterns similar to the estimates obtained for the GCD, as long as Stone-Geary parameters are allowed to be linear functions of the exogenous variables used in the GCD regressions. Also, the estimated function does not break down at higher levels of prices and provides insight into the determinants of threshold effects on water use.

2. Data and Functional Forms

2.1 Data

The data set provides information for 221 communities during the period 1981-1985.[39] Summary statistics of all variables are given in table 1. Out of 12,050 observations in the original sample, 2,620 must be dropped when years need to be identified to use the data set as a panel, table 1 also give summary statistics for the resulting 9,430 observations sample (186 communities)[40].

We use five variables that have been shown to have a significant effect on per capita water consumption (Q) across different functional forms. These five variables are the average price of water (AP), per capita income (I), percentage of the population from Spanish origin (SP), a climate variable (C), and average precipitations (AAP).[41]. The variables are described below.

Q is measured in gallons per capita per day. It is obtained from the Texas Water Development Board. The data generally excludes industrial and heavy commercial water use so

---

[39] The data set was developed by Griffin and Chang (1989). Description of the data is borrowed from their 1989, 1990, and 1991 papers. A more detailed description of the data, as well as thorough justification for data manipulations and choice of variables can be obtained from these papers.

[40] Years could only be identified when the full 5 years of a given month were reported. Consequently, the municipalities dropped were those with the most irregular reporting. The new sample is more homogeneous and is less likely to contain measurement errors. The most noticeable difference between the two samples is the drop in percentage of population of Spanish origin (see table 1).

[41] Foster and Beattie (1979) justify the exclusion of other goods in the estimation arguing that water has no direct substitutes and is a complement only with durable items such as appliances, the price of which does not affect water consumption once they have been purchased. Cross price effects are therefore normally assumed negligible in water demand.

Table 1: Summary Statistics, Samples of 12,050 and 9,430 Observations

VARIABLE	DESCRIPTION	FREQUENCY	UNITS
Q	Quantity of water produced	Monthly	Gallon/capita/day
AP	Average Price	Monthly	US$/1000 gallon
I	Income (1980 Census)	Time invariant	1000 US$/capita/year
SP	Spanish Population (1980 Census)	Time invariant	Percent of Population
C	Days with rainfall <.25 in. × Avg. temp.	Monthly	Degrees F. × days
AAP	60-year average annual precipitation	Time invariant	Inches/year

	FULL SAMPLE (years not identified)					REDUCED SAMPLE (years identified)				
	N	MEAN	ST. DEV.	MIN	MAX	N	MEAN	ST. DEV.	MIN	MAX
All	12,050					9,430				
Q		169	67.8	21.0	731		168	67.2	21.0	731
AP		1.72	0.70	0.17	6.21		1.75	0.73	0.17	6.21
I		6.40	1.68	2.30	14.5		6.50	1.69	2.30	14.5
SP		6.07	6.47	0.00	35.3		5.83	6.33	0.00	35.3
C		1801	418	644	2764		1790	412	680	2757
AAP		32.5	10.8	7.82	59.2		32.9	10.7	7.82	54.5
January	997					750				
Q		138	41.6	35.6	387		136	41.0	35.6	387
AP		1.78	0.69	0.32	6.21		1.84	0.73	0.59	6.21
I		6.39	1.67	2.30	14.5		6.49	1.62	2.30	14.5
SP		6.06	6.41	0.00	35.3		5.63	5.83	0.00	23.3
C		1287	174	871	1941		1286	161	871	1941
AAP		32.5	10.8	7.82	59.2		33.1	10.5	7.82	54.5
July	969					715				
Q		228	84.7	34.2	548		227	86.9	59.2	548
AP		1.56	0.67	0.19	5.14		1.58	0.71	0.19	5.14
I		6.44	1.69	2.30	14.5		6.64	1.76	2.93	14.5
SP		6.04	6.51	0.00	35.3		5.69	6.33	0.00	35.3
C		2359	177	1632	2764		2355	181	1632	2694
AAP		32.6	10.9	7.82	59.2		33.1	10.9	7.82	54.5

municipal water use is mostly residential. [42]

AP was shown in pretest analyses of this data (Griffin and Chang, 1990) to be the price variable on which individual decisions on water consumption are made, as opposed to a

[42] We must note, however, that the burden is on communities to report this properly and we can't be completely confident about their care. Because larger water users have larger diameter service and meters (which are accompanied by higher monthly flat fees), its easy for communities to separate this water use and report it as industrial to the TWDB.

marginal price measure suggested by economic theory.[43] The variable is calculated using the sum of water and sewer bills for all households in the community. An average household size of 2.84 is used according to Census data to get an average price in dollars per thousand gallons faced by the average household when consumption is Q.

C, the climate variable, is constructed using the number of days without significant rainfall in the community ($\geq$.25 inches) multiplied by the month's average temperature in degrees Fahrenheit. These variables were obtained from daily National Weather Service data. This variable is intended to capture the sensitivity of outdoor water demand to climate variations.

I, in thousand dollars per capita, and SP are obtained from the 1980 census. AAP is average annual precipitation in inches from 1951 to 1980, obtained from the National Oceanic and Atmospheric Administration. Finally, the dependant variable Q is obtained from the Texas Water Development Board (TWDB). Unfortunately, the TWDB obtains reports on water production rather than water consumption.[44] Q is measured in gallons per capita per day, using time invariant census data to measure population.

Q, AP, and C are time variant whereas I, SP, and AAP vary only cross-sectionally.

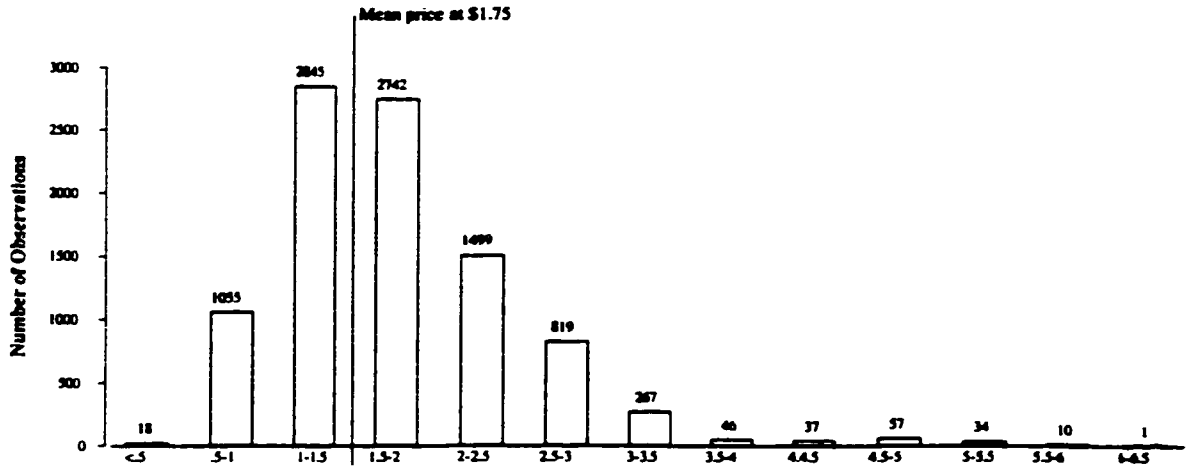
Keeping in mind our ultimate goal, which is to use these estimated functions in dynamic problems, it is important to observe the range of prices in our sample. Figure 4 shows the distribution of the price variable in the sample of 9,430 observations. The bulk of the data is within the \$1 to \$3 price range, centered around a mean price of \$1.75.

---

[43] A marginal price measure was calculated and shown to be less significant than the average price measure. This is likely due to the chosen price structure for water in Texas. However, this might not persist as water becomes more scarce.

[44] This feature is a weakness specific to water demand estimation. Two problems arise. First, the presence of storage tanks allows for monthly variations in production that may not reflect monthly variation in consumption. Some of the data had to be thrown away because of high variation that cancelled out in the next month. Second, losses to the system are included. These losses constitute 15 to 20 percent of water supply in Texas municipalities (TWDB, 1990).

Figure 4: Frequency Distribution of the Price Variable



2.2 Functional Forms

2.2.1. The Generalized Cobb-Douglas (GCD) Demand Function

The GCD specification contains features similar to a double logarithmic model in that the demand function is asymptotic to the price axis, has a positive quantity intercept and does not assume constant elasticity. The GCD form was originally proposed by Diewert (1973).[45] Its logarithmic transformation takes the following form:

$$lnQ = \delta_0 + \sum_i \sum_{j \neq i} \delta_{ij} \frac{ln(x_i + x_j)}{2} + \sum_i \delta_i ln(x_i)$$

where $\delta_{ij} = \delta_{ji}$ for all i, j and x_i, x_j are the exogenous variables (in our estimation, AP, I, SP, C, and AAP).

The GCD form is able to provide nonconstant elasticity estimates over the sample.[46]

---

[45] The GCD form has been more commonly used in production analysis. See for example Guilkey, Lovell, and Sickles (1983) for a comparison of the performance of different functional forms for cost estimation. They use an extended version of the GCD form (EGCD) which does not restrict the coefficients on cross products to 1/2. They find that in almost every comparison they conducted, the Translog (TL) systems estimator and the EGCD estimator outperform all other estimators, typically by a wide margin. However, each of these flexible forms always perform better when the production function exhibits some features of the underlying restricted form.

[46] Elasticities are calculated as $\frac{\partial lnQ}{\partial x_i} x_i$, so the elasticity of Q with respect to the variable x_i is $\delta_i + \sum_{j \neq i} \delta_{ij} \frac{x_i}{x_i + x_j}$.

However, the number of parameters involved in the functional form is large[47] and raw parameters are not readily interpretable. This complexity limits the use of estimation results to applications requiring only elasticity estimates.

2.2.2. The Stone-Geary Demand Function

Because our purpose is to use the results of our demand analysis in a structural dynamic model that requires more than just elasticity estimates, we look for a functional form that would allow price elasticity to vary seasonally and with the levels of price, income, and quantity variables, but involves fewer parameters. The functional form we propose is the Stone-Geary specification.[48] The function has been commonly used for food products, durable goods, and in the study of grants-in-aid, but has not been introduced in the water demand literature. The main advantage of the Stone-Geary specification is that it uses only two parameters and allows for nonconstant elasticities, that may increase with price. Also, both parameters have economic meaning. One of the parameters can be interpreted as a threshold below which water consumption is not affected by prices, at least not in the short-run. The magnitude of this parameter is potentially informative for investment and pricing decisions of water utilities in growing communities.

The Stone-Geary demand function is derived from a utility function of the form:

$\ln U = \sum_{i=1}^{n} \theta_i \, ln(q_i - \gamma_i)$, with $\sum_i \theta_i = 1$.

The function assumes that utilities are additive, but does not require homotheticity. Two parameters with straightforward economic interpretation highlight the function: the "mar-

---

[47] The form uses $\frac{K!}{(K-2)!2!} + K$ parameters, where K is the number of exogenous variables. In our demand estimation where 5 exogenous variables are used, the number of parameters excluding the intercept is 15.

[48] For a clear description of the theoretical features of the Stone-Geary demand function, see Deaton and Muellbauer (1980), Powell (1974), and/or Chung (1994). Stone-Geary demand and its linear expenditure system have been widely used in the study of private consumption patterns starting with the extensive study of British consumption data by Stone (1954). The system seems well suited for analyzing the demand for food products (Deaton and Muellbauer, 1980b). It has also been used in the study of public sector consumption patterns (McGuire, 1979; Johnson, 1979).

ginal budget share" (θ) and the "subsistence level" (γ).[49] The consumer is faced with a given level of income and a set of prices. The consumer first purchases a minimum acceptable level of each good (the γ_i's). The leftover income, also called "supernumerary income", is then allocated in fixed proportions to each good according to their respective preference parameter (the θ_i's).

Because we are only interested in water consumption, we can look at a utility function where the consumer derives utility from water and some aggregate of all other goods used as the numeraire. Such a utility function generates the following demand function:

$$Q_w = \gamma_w + \theta \frac{I^* - P_w\gamma_w - \gamma_z}{P_w},$$

where the subscript w indicates parameters pertaining to water and z indicates parameters pertaining to all other goods. We will use I^* for income and P for average price in this specification to distinguish it from the I and AP variables used in the GCD specification. Some minor transformations of the variables are made so that the measurement of price and income corresponds to the units used for Q (gallons per capita and per day).[50]

The price elasticity is given by:

$$\varepsilon_{ww} = -\theta \frac{I^* - \gamma_z}{P_w Q_w}.$$

The Stone-Geary function only allows for inelastic demand. This feature is not a problem for water given overwhelming support for low price elasticity in the literature.[51] Another feature of this price elasticity is that it varies inversely with price, which distinguishes it from the linear form and makes it more appropriate for expected long run elasticity changes.

The income elasticity is given by

$$\varepsilon_{wm} = \theta \frac{I^*}{P_w Q_w}.$$

---

[49] We use θ to represent the marginal budget share parameter instead of the usual β in order to keep the β variable to represent the discount rate.

[50] $I^* = I(1000/365)$, and $P = AP/1000$.

[51] See Danielson (1979) where results from a large number of studies on water demand elasticities are summarized in a convenient table.

The income elasticity implies that only normal goods can be analyzed with this functional form. Water fits into this category. A valid criticism remains that the system generates price and income elasticities that are approximately proportional to each other. This feature is difficult to justify intuitively or theoretically.[52]

Following McGuire (1979), we choose to abstract from γ_z, the general subsistence level, and treat I^* as the supernumerary income. This simplifies the function, leaving only two parameters θ and γ, both pertaining directly to water demand. We prefer to use this simpler specification since γ_z would give no information relevant to our study. Moreover, we choose to rename the remaining γ parameter as the *conditional water use threshold*. Indeed, we believe that in water demand analysis, the term "subsistence level" is misleading. The γ parameter does not indicate how much water is needed to survive, but the amount of water that is not responsive to prices. The term *conditional* is used to emphasize that this threshold is dependent on the available technology, the state of ownership, and the price of water-related durable goods during the time period of the estimation. In the following, we use the word *threshold* to refer to γ.

The parameters γ and θ are likely to vary with some exogenous variables. We assume that γ is a linear function of exogenous variables that have been shown to be of significance in water demand but are not directly included in the Stone Geary formulation. These are: the climate variable C, the Spanish population variable (SP) and the average precipitation variable (AAP). Common sense indicates that the threshold should be sensitive to these variables. The impact of climate variables is clear, the impact of the SP variable would mostly come through the fact that the variable captures income and cultural differences. The marginal budget share may also be sensitive to some exogenous variables and there is

[52] Note, however, that one finds such results in studies which are not imposing a Stone-Geary functional form. Howe and Linaweaver (1973) find an estimate of price elasticity of -0.40 and income elasticity of 0.47 for total residential demand using a logarithmic form. Gibbs (1978) finds a price elasticity of -0.51 and an income elasticity of 0.51 using a linear form and marginal price.

no reason to assume that it is constant. Actually, a constant θ would most likely generate undesirable results in terms of seasonality as the price elasticity varies inversely with quantity. Our simple formulation with other goods as the numeraire allows us to specify θ as a linear function of the same exogenous variables.[53]

3. Generalized Cobb-Douglas Estimation

3.1 Estimation Methods

We compute OLS results[54], add dummy variables for each month and year (OLSMYD), and compare the results to a random effect specification using Generalized Least Squares (GLS). The addition of time dummies is necessary to separate the seasonality and year effects from cross sectional heterogeneity.[55] The one-way (cross section) random effect model is intuitively appealing in this setup since the sample can be considered to be a random sample from a larger population of communities. Two features of the data make a fixed effect model impractical. First, the number of cross sections (N) is large and the loss of degrees of freedom caused by the addition of N-1 dummy variables jeopardizes the stability of the coefficients. Second, and most important in our case, three of the variables used are time invariant, which makes a fixed effect model incapable of identifying coefficients on these variables.

The one-way random effect specification of the GCD form is

$$lnQ_{it} = \delta_0 + Z_{it}\theta + u_i + \varepsilon_{it},$$

where Z is a vector of right hand side variables to be used in the GCD form, ε_{it} is the clas-

---

[53] Johnson (1979) also assumes that γ is a linear function of some exogenous variables. He argues that there is no theoretical reason to assume that θ is not also function of exogenous variables, but in his analysis with more than one θ, it would have seriously complicated the analysis, because of the restriction that the θ's must add up to one.

[54] The OLS results reproduce previously published results by Griffin and Chang, 1991.

[55] We started with adding only month dummies for seasonality but we found that year effects were highly significant. A simple OLS linear regression of Q on the exogenous variables, month dummies, and a year variable gave a highly significant coefficient of 4.82 on the year variable. Adding year dummies to our GCD regression takes care of this positive trend.

sical error term, normally distributed and with zero mean, and u_i is a community specific error term. This panel data formulation of the model uses the heterogeneity information contained in the data. The inclusion of a community-specific error term generates a non-standard covariance structure. If the u_i's are non-zero, GLS is efficient and improves upon OLS as long as the effects are uncorrelated with the regressors, which would make GLS inconsistent.[56] Hausman-Wu tests are carried out for all the random effect estimations. The results for all one-way random effect specifications suggest that there is no problem of correlation between the effects and the exogenous variables when only cross section random effects are used. We also report F-tests from the fixed effects regression (table 2), all indicating the presence of sample heterogeneity among cross section. The GLS one-way random effect estimation should therefore be efficient and improve upon OLS.

3.2 GCD Results

We perform the OLS and GLS regression with and without time dummies.[57] Our regression results in a raw form (i.e., with coefficients and standard errors pertaining to the regressors) for OLS and GLS estimations are reported in table 2. Since regressors in the GCD form are composites of our exogenous variables, raw results are not readily interpretable. To quantify the impact and significance of each exogenous variable separately, we calculate elasticities and their corresponding standard errors.[58] Results are reported in table 3.

To evaluate the effect of the change in sample size from 12,050 to 9,430 observations,

---

[56] For a theoretical discussion of one-way fixed and random effects models, see Greene (1997), pp. 618-630 and Baltagi (1995), pp. 9-18.

[57] Estimation of the GLS one-way unbalanced data random effects model is performed in SAS-ETS (1996) using the Baltagi and Chang (1994) specialization of the approach used by Wansbeek and Kapteyn (1989) for unbalanced two-way models. We also carry out the two-way random effect estimation, but the Hausman-Wu test on this regression gives a poor result suggesting some correlation between the time effects and the exogenous variables (see table 2).

[58] The price elasticity is calculated as:

$$\varepsilon_{ww} = AP * \left[\frac{\hat{b}[\ln(AP)]}{AP} + \frac{\hat{b}[\ln(AP+I)]}{AP+I} + \frac{\hat{b}[\ln(AP+SP)]}{AP+SP} + \frac{\hat{b}[\ln(AP+C)]}{AP+C} + \frac{\hat{b}[\ln(AP+AAP)]}{AP+AAP} \right]$$

where $\hat{b}[.]$ denotes the estimated coefficient on [.].

Table 2: Generalized Cobb-Douglas Regressions

Regression Type	OLS[1]	OLS	OLS	GLS1[2]	GLS1	GLS1	GLS2[2]
Sample size	12050	12050	9430	9430	9430	9430	9430
Dummies	-	Month/Year	Month/Year	-	Month	Month/Year	-
Intercept[3]	-5.08**	-1.72**	-2.80**	-6.31	-2.84	-2.12	-1.95
	(0.39)	(0.34)	(0.39)	(1.88)	(2.02)	(2.95)	(2.77)
ln(AP)	-0.58**	-0.60**	-0.60**	-0.41	-0.40	-0.51	-0.50
	(0.026)	(0.025)	(0.026)	(0.037)	(0.036)	(0.035)	(0.035)
ln(AP+I)	0.48**	0.82**	0.58**	-0.92	-0.74	-0.87	-0.78
	(0.15)	(0.14)	(0.15)	(0.20)	(0.19)	(0.19)	(0.18)
ln(AP+SP)	0.12**	0.10**	0.05**	0.058	0.09	0.13	0.13
	(0.016)	(0.015)	(0.016)	(0.035)	(0.034)	(0.033)	(0.032)
ln(AP+C)	124**	57.5**	42.6*	164	108	86.6	80.0
	(18.9)	(18.7)	(20.2)	(16.74)	(16.48)	(16.0)	(15.78)
ln(AP+AAP)	-0.58	-0.30	1.40**	0.93	1.17	2.24	2.06
	(0.36)	(0.35)	(0.39)	(0.53)	(0.51)	(0.50)	(0.49)
ln(I)	-0.10	-0.34**	-0.34**	1.03	1.01	1.23	1.21
	(0.11)	(0.11)	(0.12)	(0.71)	(0.76)	(1.11)	(1.04)
ln(I+SP)	-0.36**	-0.27**	-0.063	-0.17	-0.18	-0.24	-0.27
	(0.044)	(0.04)	(0.048)	(0.46)	(0.50)	(0.73)	(0.68)
ln(I+C)	71.3**	35.0**	50.2**	37.12	-11.9	-23.7	-40.0
	(7.96)	(7.96)	(9.05)	(7.92)	(8.02)	(7.78)	(7.71)
ln(I+AAP)	-0.96**	-0.74**	-0.91**	-1.08	-0.78	-0.82	-0.67
	(0.19)	(0.18)	(0.20)	(3.27)	(3.51)	(5.17)	(4.85)
SP	0.010**	0.012**	0.005	0.0071	0.010	0.012	0.016
	(0.0030)	(0.0030)	(0.0034)	(0.049)	(0.052)	(0.077)	(0.072)
ln(SP+C)	18.0**	2.44	5.24	32.8	19.4	16.4	9.95
	(2.92)	(2.95)	(3.33)	(2.44)	(2.47)	(2.40)	(2.39)
ln(SP+AAP)	-0.54**	-0.40**	-0.56**	-1.00	-0.92	-0.91	-0.85
	(0.061)	(0.060)	(0.066)	(0.99)	(1.06)	(1.56)	(1.46)
ln(C)	-245**	-118**	-127**	-284	-160	-122	-88.7
	(19.7)	(20.1)	(22.1)	(17.14)	(17.66)	(17.2)	(17.3)
ln(C+AAP)	32.3**	24.5**	30.2**	52.48	46.44	44.4	40.0
	(1.70)	(1.71)	(2.00)	(1.51)	(1.52)	(1.47)	(1.47)
ln(AAP)	1.08**	0.67*	-0.81*	-0.24	-0.67	-1.59	-1.53
	(0.35)	(0.34)	(0.38)	(2.44)	(2.61)	(3.81)	(3.58)
R2[4]	0.44	0.47	0.48	0.57	0.60	0.62	0.33
F test	633	355	294	--	--	--	--
Hausman-Wu Test (m)	--	--	--	0.23	0.23	0.096	14.1
DF for Hausman-Wu test				9	20	24	9
(prob>m)	--	--	--	(.99)	(.99)	(.99)	(0.12)

** significant at the 1 percent level or better, *significant at the 10 percent level or better

(1) Benchmark OLS : Replica of results in Griffin and Chang (1991) (2) GLS1 = One-way random effects, GLS2 = two-way random effects

(3) When dummies are included intercept corresponds to month of December, year 1985

(4) Adjusted R2 are reported for OLS regressions. For GLS regressions, a generalized version of R2 is used.

we also report the OLS results with and without time dummies for both samples. Finally, to investigate how the function performs in picking up seasonal variation in price elasticity,

we calculate the parameters and elasticities at the January and July means. Table 4 presents these results.

Table 3: Results from GCD Estimations

Elasticities and Standard Errors Evaluated at the Sample Means

	OLS	OLS	OLSMYD	OLSMYD	GLS	GLSMYD
N	12,050	9,430	12,050	9,430	9,430	9,430
AP	-0.37**	-0.35**	-0.37**	-0.36**	-0.39**	-0.47**
	(0.004)	(0.005)	(0.004)	(0.005)	(0.007)	(0.007)
I	0.19**	0.12**	0.17**	0.11**	0.17	0.19
	(0.002)	(0.002)	(0.002)	(0.002)	(0.034)	(0.054)
SP	-0.04**	-0.04*	-0.03**	-0.03**	-0.04	-0.03
	(0.0009)	(0.0005)	(0.0003)	(0.0008)	(0.011)	(0.018)
C	0.72**	0.72**	0.53**	0.56**	0.74**	0.53**
	(→ 0)	(0.00001)	(0.00001)	(→ 0)	(→ 0)	(0.00002)
AAP	-0.15**	-0.18**	-0.14**	-0.17**	-0.16	-0.12
	(0.0003)	(0.0004)	(→ 0)	(0.0004)	(0.006)	(0.009)

** Significant at the .01 level or better. *significant at the .1 level or better.
(→ 0): Standard error $< 10^{-5}$

Table 4: Results from GCD Estimations : Seasonality

Elasticities and Standard Errors Evaluated at January and July Means

	JANUARY			JULY		
	OLS	OLSMYD	GLSMYD	OLS	OLSMYD	GLSMYD
N	12,050	9,430	9,430	12,050	9,430	9,430
AP	-0.31**	-0.33**	-0.43**	-0.41**	-0.39**	-0.49**
	(0.006)	(0.007)	(0.008)	(0.006)	(0.006)	(0.008)
I	0.28**	0.18**	0.16	0.14**	0.08**	0.19
	(0.003)	(0.003)	(0.05)	(0.002)	(0.002)	(0.05)
SP	-0.02**	-0.02**	-0.01	-0.05**	-0.03**	-0.04
	(0.001)	(0.001)	(0.02)	(0.001)	(0.001)	(0.02)
C	0.32**	0.25**	0.19**	0.96**	0.74**	0.74**
	(→ 0)	(0.00001)	(0.00002)	(→ 0)	(0.00001)	(0.00001)
AAP	0.07*	0.03*	0.18	-0.29**	-0.29**	-0.30
	(0.0005)	(0.0005)	(0.009)	(0.0004)	(0.0004)	(0.009)

** Significant at the .01 level or better. * Significant at the .05 level or better.
(→ 0): Standard error $< 10^{-5}$

Price elasticities obtained with OLS and evaluated at the sample mean range from 0.35 to 0.37. The GLS one-way random effect estimates of price elasticities are uniformly larger than their OLS counterparts, ranging from 0.39 to 0.47. As for the OLS estimate, the GLS estimate is a weighted average of the "within" estimate (which ignores variation between

communities) and the "between" estimate (which ignores variation within communities). However, a non-zero variance of the u_i's increases the weight on the within estimate and decreases the weight on the between estimate. Since the price (AP) and the climate variable (C) are the only variables that have a non-zero within estimate, we expect the GLS estimates on these variables to be greater. In particular, since income is time invariant, the within estimate of price elasticity is likely to include the missing variation in income (which should increase price elasticity since water is a normal good). Finally, the use of GLS affects the significance levels of our estimates. Although the significance level on price elasticities remain high, significance levels on elasticities with respect to I, SP, and AAP are reduced. Again, since the within variation is given more weight in the GLS regression and since these variables are invariant within communities (and therefore given a zero coefficient in the within estimation), this is not a surprising result.

We need to evaluate how elasticities are likely to behave if this functional form and our estimates are used out of sample in long-run planning models (i.e., when higher values of variables such as price and income are used). To do so, we calculate the elasticities at the maximum sample values of price and income. Results are reported in Table 5. Elasticities are generally reduced except when evaluated at maximum income.

Table 5: Behavior of Price Elasticity as Price, Income, and Quantity increase

Estimation	OLS	GLS	OLSMYD	GLSMYD
At maximum Price	0.13	-0.11	0.08	-0.21
At maximum Income	-0.37	-0.29	-0.42	-0.38
At maximum Price and Income	0.08	0.06	-0.03	-0.05

Note: Elasticities are evaluated at sample averages, unless maximums are used.

One important feature is that in most cases, except for the random effect specification with dummies, price elasticities become positive. This indicates that the functions do not perform as well at the maximums. This problem is likely to worsen when the estimates

are used out of sample. Figures 5 and 6, provide a comparison between the OLS and GLS estimations of the GCD function with and without month and year dummies. As expected, the GCD function estimated with random effects (GLS) is better behaved than the OLS estimated function at higher prices, retaining its downward sloping shape. Figure 7 and 8, provide a comparison of the January and July functions using OLS and GLS with and without the use of month and year dummies. Again, the GLS estimated function retains its negative slope. The seasonal shifts are somewhat less pronounced at higher prices when using GLS. Finally, the use of dummy variables in the estimation seems to capture the shifts in a smoother way over a wider price range.

Figure 5: Plot of Estimated GCD Function

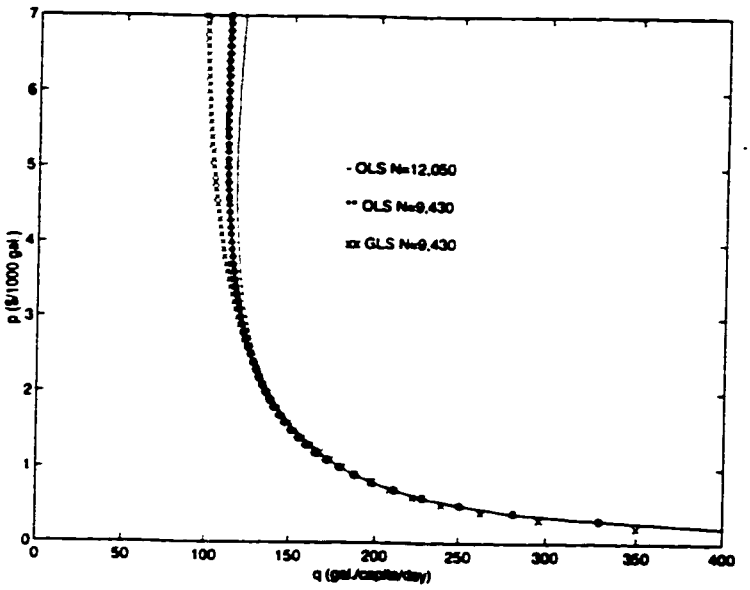


Figure 6: Plot of Estimated GCD Function with Month and Year Dummies

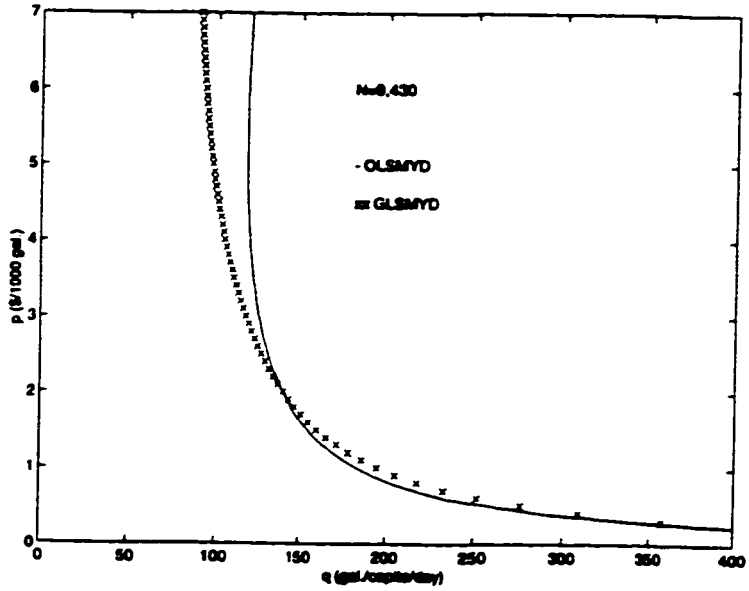


Figure 7: Plot of Estimated GCD Function: Seasonality

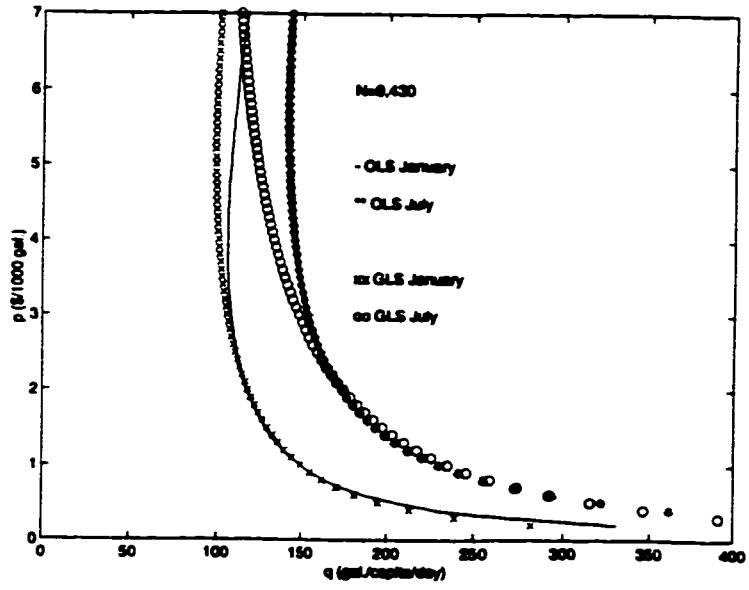
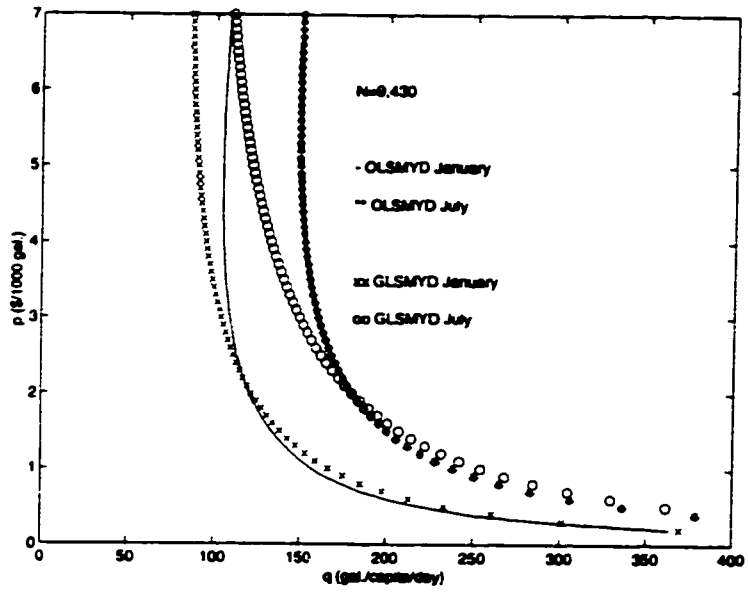


Figure 8: Plot of Estimated GCD Function with Month and Year Dummies: Seasonality



4. Stone-Geary estimation

We perform regressions for the Stone-Geary functional form using OLS and GLS (one-way random effects) estimators. The form can be written as linear or nonlinear in the parameters. Although both forms are nonlinear in the variables, we refer to these two estimation techniques as linear and nonlinear. The linear estimation focuses on the intercept and θ, which are estimated as independent coefficients, γ is then calculated outside of the estimation. The nonlinear estimation estimates both θ and γ simultaneously.. Since the linear estimation is likely to lack efficiency, we use nonlinear OLS and nonlinear GLS to estimate the parameters. Nonlinear results are found to be very similar to the linear results (only two iterations are needed to complete the estimation), but the nonlinear results allow us to calculate and report approximate standard errors and significance levels on the estimates of γ and θ, and therefore, on price elasticity. We use the linear regressions for Hausman-Wu tests. The results on the threshold parameter and elasticities are discussed below. Results

on elasticities and the shape of the estimated demand functions are compared to the GCD specification found best in the previous section, namely the random effect specification with month and year dummies.[59]

4.1 Estimated Equations

SGE (γ) refers to the estimation where γ is a linear combinations of the exogenous variables but not θ. The estimable Stone-Geary equation with fixed θ takes the form:

$Q = (1-\theta)(\alpha_o + \alpha_1 C + \alpha_2 SP + \alpha_3 AAP) + \theta\frac{I^*}{p}$,

where I^* and p are as defined in section 2 and $\gamma = \alpha_o + \alpha_1 C + \alpha_2 SP + \alpha_3 AAP$.

For the linear estimation, the estimable equation is:

$Q = \alpha'_o + \alpha'_1 C + \alpha'_2 SP + \alpha'_3 AAP + \theta\frac{I^*}{P}$,

where $\alpha'_i = \frac{\alpha_i}{1-\theta}$. Using the estimates of θ and α'_i, one can calculate γ.

SGE (γ, θ) refers to the estimation where both γ and θ are linear combinations of the exogenous variables. The Stone-Geary equation with variable θ takes the form:

$Q = [1 - (\theta_o + \theta_1 C + \theta_2 SP + \theta_3 AAP)](\alpha_o + \alpha_1 C + \alpha_2 SP + \alpha_3 AAP) + (\theta_o + \theta_1 C + \theta_2 SP + \theta_3 AAP)\frac{I^*}{P}$.

where γ is defined as above, and $\theta = \theta_o + \theta_1\overline{C} + \theta_2\overline{SP} + \theta_3\overline{AAP}$,where $\overline{X}$ denotes the sample mean of variable X.

For the linear estimation, the estimable equation is:

$Q = \alpha'_o + \alpha'_1 C + \alpha'_2 SP + \alpha'_3 AAP + \theta_o\frac{I^*}{P} + \theta_1 C\frac{I^*}{P} + \theta_2 SP\frac{I^*}{P} + \theta_3 AAP\frac{I^*}{P}$,

where $\alpha'_i = \frac{\alpha_i}{1-\theta}$ and $\theta = \theta_o + \theta_1\overline{C} + \theta_2\overline{SP} + \theta_3\overline{AAP}$.

---

[59] For both the linear and non linear estimations, an error term is added to the regression, which we omit in writing the following estimable equations. The error term takes the structure ϵ_{it} in the OLS regressions and $u_i + \epsilon_{it}$ in the GLS specifications, where ϵ_{it} and u_i have the usual classical assumptions. A random utility assumption would cause the ϵ_{it} to be heteroskedastic (Brown and Walker, 1989), however, we find no evidence of this being a significant problem in the regressions.

4.2 Transformation of Variables for Nonlinear GLS

In order to estimate the non-linear Stone-Geary form with non-linear GLS, we first transform the variables.[60] We decompose the error variance into two components: σ_ϵ^2, the variance specific to the error, and σ_u^2, the variance specific to the cross sectional effects. The panel is unbalanced: each cross section i has Ti observations and $\sum_i T_i = N$. There are K regressors excluding the intercept. We perform a variance decomposition due to Greene (1997):

$m_{ee} = \frac{e'e}{\sum_i T_i}$, for $i = 1$ to N;

m_{ee} is the mean square error from the linear OLS regression. We use OLS estimates rather than within estimates from a least square dummy variable estimation because of the presence of time invariant variables. Next we use

$m_{**} = \frac{e_{**}'e_{**}}{N-K}$,

$e_{**} = Q - \widehat{\beta X}$

where the $\widehat{\beta}$ are the OLS parameter estimates and $\overline{X}$ is a vector of means of exogenous variables. We use the probability limits:

$m_{ee} = \sigma_\epsilon^2 + \sigma_u^2$ and

$\sigma_u^2 = m_{**} - \sigma_\epsilon^2 * \widehat{Q}_n$, where $\widehat{Q}_n = 1/N * \sum_{i=1}^{N} 1/T_i$.

From these equations we can solve for the estimated $\widehat{\sigma}_\epsilon^2$ and $\widehat{\sigma}_u^2$ using the following:

$\widehat{\sigma}_\epsilon^2 = \frac{1}{1-\widehat{Q}_n} * (m_{ee} - m_{**})$

$\widehat{\sigma}_u^2 = \frac{1}{1-\widehat{Q}_n} * m_{**} - \frac{\widehat{Q}_n}{1-\widehat{Q}_n} * m_{ee}$

All the variables in the model (including intercept and dummies) are transformed using:

$Xt_i = X_i - \beta_i \overline{X}$,

$\beta_i = 1 - \frac{\widehat{\sigma}_\epsilon}{\sqrt{Ti * \widehat{\sigma}_u^2 + \widehat{\sigma}_\epsilon^2}}$.

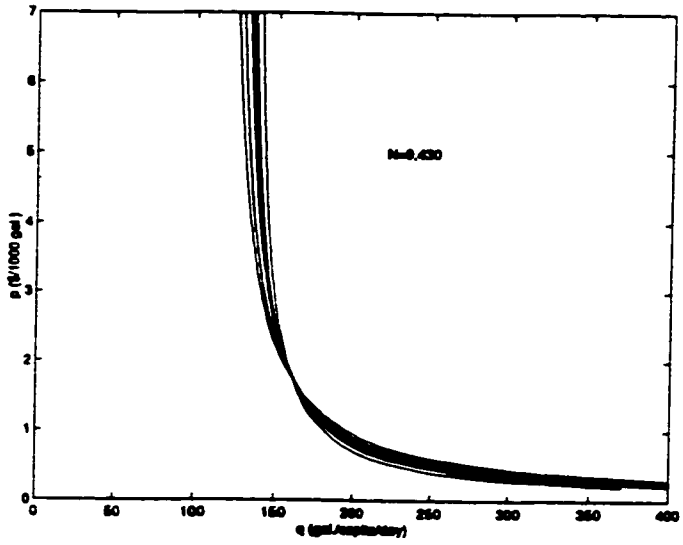
---

[60] Available statistical packages handles linear GLS for estimating a random effect estimation but do not handle nonlinear GLS. Variables need to be transformed first to get the desired results with the non-linear OLS estimator.

5. Stone-Geary Results and Comparison to GCD

Raw results from all different specifications and estimation procedures are provided in table 8 (linear results) and table 9 (nonlinear results) at the end of the chapter (pp. 66 and 67. To evaluate the robustness of our results, we also plot the estimated functions evaluated at the sample means for all the different specifications using the nonlinear estimates evaluated at the sample means (figure 9). The graph shows that the variation between the specifications is so small that it is difficult to distinguish between them.

Figure 9: Plots of all Nonlinear Estimations of Stone-Geary Specifications Evaluated at the Sample Means



Hausman-Wu tests for GLS random-effect regressions are quite poor when month and year dummies are not included. These test results suggest there is some correlation between the regressors and the effects, and cast doubt on consistency of the parameter estimates.[61] Since the Hausman-Wu test results for the GLS regressions with month and year dummies are significantly better, we focus on these results. The hypothesis of correlation is rejected

[61] Since nonlinear and linear results are very similar we do not recalculate the Hausman-Wu tests for nonlinear regressions. Hausman-Wu tests on linear regressions are reported in table 8 at the end of the chapter.

at the 0.01 level for SGE(γ). Hausman-Wu test results are not as good when θ is variable, and we cannot reject the hypothesis of correlation at the 0.1 level. When θ is not fixed, the regression relies more heavily on variables that are combined with time invariant exogenous variables. The Hausman-Wu test drops the purely time invariant variables that enter in the estimation of γ, but it does leave all the time invariant variables in θ, since they are all divided by P, which is time variant. Since the fixed effect estimator on these variables must be zero the difference between the fixed effect estimator and the random effect estimator is likely to increase, without necessarily implying correlation. This argument casts some doubt on the straightforward interpretation of the Hausman-Wu test when time invariant variables are present and combined with other variables in a non-additive way. The addition of dummies might improve the test not so much by taking care of the correlation problem but by reducing the reliance on time invariant variables. However, since the test remains informative when it turns out positive, we still prefer to rely on the regressions with the better Hausman-Wu test results.

Table 6 reports the estimated marginal budget shares (θ), conditional water thresholds (γ), and price elasticities (ε_{ww}) for the SGE(γ) and SGE(γ, θ) regressions with dummies. The parameters and associated approximate standard errors are evaluated at the sample means, the January means, and the July means.[62] All parameters are highly significant.

Conditional water threshold: Results on the average threshold parameter in gallons per capita per day (gcd) range from 114 to 128 depending on whether it is calculated at annual means or at selected monthly means. The January threshold is estimated between 86 and 118 gcd. The July threshold jumps to 155-176 gcd. Seasonal variations in γ are reduced when θ also picks up some of the variation (when θ is a function of exogenous

[62] The threshold parameter involves an intercept which includes month and year dummies. For results evaluated at sample averages, the averages of month dummies and year dummies are used. For the January and July results, γ is calculated using the estimated coefficient on the appropriate month dummy.

Table 6: Nonlinear Estimation of Stone-Geary with Month and Year Dummies

Conditional Water Thresholds (γ),
Marginal Budget Shares (θ) and Price Elasticities (ε_{ww})

Specification		SGE (γ)		SGE (γ, θ)	
Estimator		OLS	GLS	OLS	GLS
All Seasons	γ	123	114	125	128
		(5.16)	(12.6)	(8.10)	(14.61)
	θ	0.0038	0.0046	0.0036	0.0032
		(9.3E-5)	(0.00014)	(0.00050)	(0.00046)
	ε_{ww}	-0.23	-0.28	-0.22	-0.19
		(0.0057)	(0.0082)	(0.029)	(0.028)
January	γ	95	86	115	118
		(5.08)	(12.5)	(7.77)	(14.52)
	θ	0.0038	0.0046	0.0021	0.0017
		(9.3E-5)	(0.00013)	(0.00048)	(0.00047)
	ε_{ww}	-0.27	-0.33	-0.15	-0.12
		(0.0057)	(0.0082)	(0.029)	(0.029)
July	γ	176	166	155	162
		(6.76)	(13.1)	(9.54)	(15.17)
	θ	0.0038	0.0046	0.0053	0.0047
		(9.3E-5)	(0.00013)	(0.00055)	(0.00048)
	ε_{ww}	-0.19	-0.24	-0.27	-0.24
		(0.0057)	(0.0082)	(0.034)	(0.029)

Note: All parameters are significant with a confidence level >0.99

variables).

We need to be very careful in interpreting γ for several reasons. First of all, water production data is used instead of water consumption. This implies that all losses to the system are included in the threshold. Water utilities (or city planners) need to interpret this parameter as a per capita production threshold, which can be lowered proportionately by a reduction in losses. To obtain the consumer's actual water use threshold, the γ estimate must be lowered to reflect average system losses in the community.[63]

[63] The Texas Water Development Board estimated losses to the system to be between 15 and 20 percent of water production (TWDB, 1990). Taking an average loss of 20 percent, an approximate water consumption function can be recovered by reducing total quantity and all raw coefficient estimates by 20 percent. Since both θ and Q must be reduced by 20 percent, results on elasticities remain unchanged.

The second caveat is that all other goods are assumed to be gross complements with water consumption. In particular, the trade-off between a more expensive water-efficient durable good versus one that is less expensive but uses more water is not considered. Threshold parameters are likely to be reduced if technology changes allow water use to be economized. The availability of water-efficient durable goods, such as water-efficient toilets and washing machines,will have an impact on residential demand and the water use threshold. However, this will be observed only if during the sample period there is a large price reduction in these goods. Also, residential customers are likely to reassess their decision to maintain lawns if the price of water increases beyond the levels of our sample, a fact that cannot be captured in our estimates.

Finally, it is important to note that the estimated threshold is conditional on the price structure in place. A price structure that incorporates fixed charges does not allow marginal prices to be clearly defined. Consequently, individuals use average price in their decision to consume water, as is the case in the Texas sample. A large threshold indicates that conservation practices which are not price based (or only affect prices of substitutes) may potentially have more impact on the reduction of water use per capita than price increases. This does not mean that a fundamental change in the price structure, making marginal prices more transparent, would not be a better policy tool.

These considerations cast make our results problematic for long-run predictions when increased water scarcity would bring about modifications of the pricing system and change the value of γ. In addition, the water threshold is only a function of our climate variable, average rainfall, and percent of population of Spanish origin. Only the Spanish origin variable is likely to increase and have some impact on the level of γ.[64] Ideally, if we had a longer

[64] SP impacts the threshold negatively and the elasticity positively indicating that the Mexican population is more flexible in their water use. γ is significantly reduced in all the estimations when the maximum value of SP (35.3 percent) is used. When elasticities are allowed to vary with SP, we obtain values of γ ranging from 70 to 88 gcd (compared to 124-134 gcd whith the mean value of SP). However, we would expect the coefficient on SP to be reduced over time if the Mexican population becomes more integrated.

time series and data on the price and availability of water conserving goods, we would make γ a function of these new variables as well, allowing γ to vary over time. Nonetheless, even when production data is used, results on this conditional threshold parameter are still informative for the short-run. They show given the available range and structure of water prices, the availability and prices of other goods that could partially "substitute" for water, and with losses of 15-20 percent, more than half of water use is unresponsive to prices.

Elasticities: Price elasticity is calculated as $\varepsilon_{ww} = -\theta(.)\frac{I^*}{P.Q}$, evaluated at the appropriate means. The elasticities obtained from the Stone-Geary specification are lower than the GCD results. The range of elasticities using Stone-Geary is .19 to .28, compared to .35 to .47 using GCD. The results also indicate that the Stone-Geary specification allows significant seasonal variation in elasticities. When θ is fixed, summer elasticities are lower than winter elasticities. This is not surprising since the elasticity is inversely proportional to the expenditure; with θ fixed, the increase in Q, due mostly to factors other than a price decrease, dominates. If one wants to measure seasonal elasticities, the marginal budget share parameter needs to vary with the seasonal variables (C, in our estimation). When θ is allowed to vary, the result of higher price elasticities for summer demand is re-established. Both the OLS and GLS estimation of SGE(γ, θ) give a seasonal variation of elasticity estimates larger than the GCD regression. Stone-Geary results give an increase in elasticity between winter and summer of .12 (from .15 to .27 for OLS and between .12 and .24 for GLS), whereas the GCD results showed an increase of .6.

Our main goal is to determine whether the simple Stone-Geary specification could give us ranges of elasticities comparable to the more flexible GCD form, yet allow for out-of-sample use in dynamic simulations when variables increase over time. Table 7 reports price elasticities at the sample maxima of price, income, and quantity.

It shows that the Stone-Geary specification yields a large variation in price elasticity

Table 7: Behavior of Price Elasticity as Price, Income, and Quantity Increase

$$\varepsilon_{ww} = -\theta(.)\frac{I^{\bullet}}{P.Q}^{(1)}$$

Specification	SGE (γ)		SGE (γ, θ)	
Estimator	OLS	GLS	OLS	GLS
At maximum Price	-0.065	-0.079	-0.062	-0.055
At maximum Price and Quantity	-0.015	-0.018	-0.014	-0.013
At maximum Income	-0.51	-0.63	-0.49	-0.43
At maximum Income and Quantity	-0.12	-0.14	-0.11	-0.10
At maximum Price, Income, and Quantity	-0.033	-0.041	-0.032	-0.028

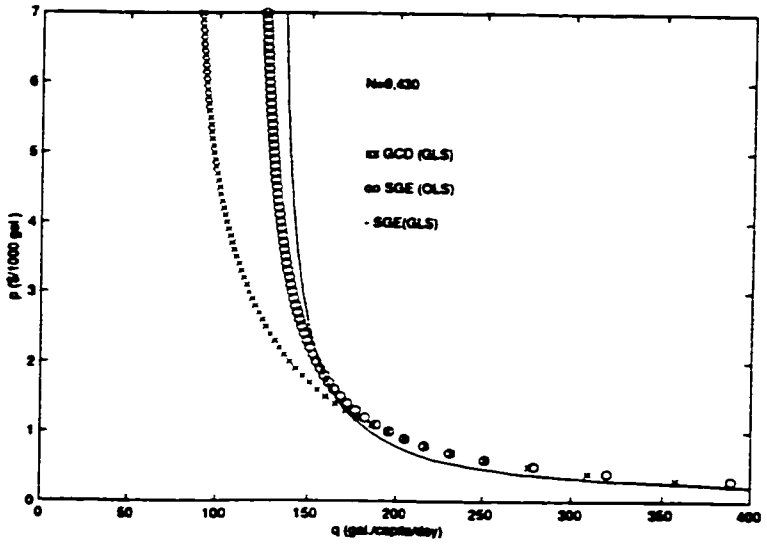
(1) Evaluated at sample averages, unless maximums are used as indicated.

when the values of most variables increase. All elasticities calculated at maxima instead of averages are lower than elasticities calculated at the sample averages. The only exception is when the elasticity is calculated at maximum income only, which seems reasonable since the price relative to income has decreased. Price elasticities remain strictly negative in all cases, which was not true for the GCD estimation.[65]

Figure 10 shows the plot of the OLS and GLS estimated Stone-Geary function with month and year dummies and nonconstant marginal budget shares [SGE(γ, θ)] and compares it to our previous results using GCD. We use the one-way random effect estimated GCD with month and year dummies, which was shown to behave best at higher prices. The Stone-Geary estimated functions have a steeper slope and retain higher quantity levels than GCD at higher prices. This is a result of specifying a threshold parameter which constrains the function to be asymptotic to a minimum quantity. The function is still well behaved in that it is downward sloping everywhere, which was not the case with the OLS estimated GCD function (figure 6).

[65] Compare these results to table 5.

Figure 10: Comparison of estimated GCD and Stone-Geary



6. Conclusion

Our motivation for using Stone-Geary is simple: we wanted a function with few, easily interpretable, parameters that could be used in dynamic management problems that require more than just price elasticities on the demand side. Alternative functional forms that retain a simple parameter structure are the linear and log-linear forms. However, linear demand forces the elasticity to increase as price increases, and the log-linear form maintains constant price elasticity. Constant price elasticity has been commonly used in water management models that include the computation of consumer surplus. However, research using more flexible functional forms provides strong evidence that water demand elasticities vary seasonally and over time.

It is surprising that the Stone-Geary function has not been used in water demand estimation. The behavior it implies, that individuals will use a minimum amount of a good disregarding price fluctuations, and then decide on the rest according to a preference parameter, the price, and their income, seems well suited for water. Furthermore, the fact

that water is a normal good and has inelastic demand makes it a good candidate for a Stone-Geary specification. The form is parsimonious enough to parameterize demand in structural dynamic models of water management and our analysis shows that the specification is also able to provide estimates of elasticities comparable to other more complicated forms.

We needed a good benchmark to evaluate the performance of the Stone-Geary specification. We refined the General Cobb-Douglas (GCD) estimation of municipal water demand in Texas using econometric techniques that allow us to fully use the panel structure of our data. We found that the new random effect estimation (GLS) improves upon the OLS procedure. Price elasticities are slightly higher but remain quite similar and the function remains well behaved at higher prices. We obtained elasticity estimates of 0.39 to 0.47 using GLS, compared to 0.37 using OLS. Summer elasticities were on average .8 higher than winter elasticities. GCD gives price elasticities that decrease both in price and income. When calculated at the maxima of our price and income data, elasticities were significantly lowered. Of all our estimations of GCD, only GLS with month and year dummies retained a negative price quantity relationship at high prices.

Our Stone-Geary results give elasticities ranging from 0.19 to 0.28. Although lower, and likely not providing as good a fit within the sample as the GCD results, these estimates are not out of the range of price elasticity estimates in the literature. Summer elasticities were .12 higher than winter elasticities, indicating that Stone-Geary can successfully estimate the expected seasonal variation. It is important to note that the marginal budget share parameter θ in Stone-Geary must be allowed to vary with the monthly climate variable to get this result. Stone-Geary gives estimates of price elasticities that increase with income but decrease with price and quantity. At the maximum price, income, and quantity we obtained elasticities in the 0.03 to 0.04 range. These results are similar to the ones obtained with GCD. Such results cannot be obtained when using linear or log-linear forms. Moreover, the Stone-Geary

estimation results for these elasticities are more robust to the type of estimator than the GCD results, and the negative price quantity relationship is preserved. This is encouraging if we want to use the estimation results over a larger range of prices.

Finally, we look at the Stone-Geary results on the conditional water use threshold. We find that approximately three quarters of the total water usage is not responsive to price changes. This threshold needs to be lowered to reflect the 15-20 percent system losses in Texas communities that creates a gap between water production and water consumption. However, this still leaves more than half of water demand unresponsive to price increases. We cautioned the reader about the conditional nature of this parameter. Ideally we would want to include some long term determinants of the threshold water use. Indeed, increased water prices are expected to eventually affect the decision to buy durable goods that require the use of water, which in turn will reduce the water use threshold. Also, increased water scarcity is likely to bring about price reforms to give the right price signals to consumers. This is why a longer term analysis would be necessary to look at the behavior of this threshold over time. Our estimated water threshold, nonetheless, gives information conditional on the type of data used and on the length of the time series. For time periods limited to five years and given a price range of about 0.2 to 6.2 dollars per thousand gallons provided by our data, water consumption cannot on average be reduced by more than 40 percent using exclusively a price instrument of the current type.

The exclusion of all other prices and, in particular, prices of durable goods that need water, result in an important reservation about using the Stone-Geary results in long-run management problems. More data is needed to include these variables in the analysis. They could be included without changing the structure of the Stone-Geary system by simply adding more exogenous variables as determinants of the threshold level and possibly the budget share parameter.

Further work could compare the result of the Stone-Geary to a nonparametric or semi-parametric demand estimation such as in Hausman and Newey (1995). Such types of estimations have not yet been used in the water demand literature but could reduce the possibility of mispecification, provide an even less restrictive benchmark than the GCD, and allow us to look more closely at the size and variation in elasticities.

For the purpose of this dissertation, we believe that we have shown that the Stone-Geary function has the required characteristics to be a good functional form to use in our dynamic program of water management. The consumer surplus from the Stone-Geary is defined as long as we use a fixed part determined by the threshold water use and a functional part for any water use above this level. We give more details on the form of the benefit function in the next chapter.

Table 8: Stone-Geary: Linear Regressions Results

$$Q = \alpha_o' + \alpha_1' C + \alpha_2' SP + \alpha_3' AAP + \theta \frac{I_r^*}{P}^{(1)}$$
$$\alpha_i' = \frac{\alpha_i}{1-\theta}, \theta \text{ is } \theta \text{ or } \theta_o + \theta_1 C + \theta_2 SP + \theta_3 AAP$$

Regression Type	OLS	GLS	OLS	GLS	OLS	GLS	OLS	GLS
Sample Size	9430	9430	9430	9430	9430	9430	9430	9430
$\theta = f(C, SP, AAP)$	No	No	No	No	Yes	Yes	Yes	Yes
Month and Year Dummies	No	No	Yes	Yes	No	No	Yes	Yes
α_0' [2]	39.72**	28.72**	59.66**	43.70**	109.25**	32.09**	117.02**	33.60*
(standard error)	(3.36)	(11.88)	(4.78)	(12.24)	(7.39)	(16.08)	(8.05)	(15.98)
α_1'	0.075**	0.076**	0.070**	0.071**	0.034**	0.036**	0.035**	0.037**
	(0.0013)	(0.0010)	(0.0028)	(0.0024)	(0.0029)	(0.0022)	(0.0038)	(0.0031)
α_2'	-0.59**	-0.55	-0.54**	-0.46	-1.82**	-1.57**	-1.84**	-1.54**
	(0.097)	(0.49)	(0.095)	(0.50)	(0.22)	(0.62)	(0.21)	(0.61)
α_3'	-1.46**	-1.36**	-1.43**	-1.29**	-1.01**	1.41**	-1.02**	1.44**
	(0.058)	(0.30)	(0.056)	(0.30)	(0.13)	(0.40)	(0.13)	(0.40)
θ or θ_0	0.0037**	0.0042**	0.0038**	0.0046**	-0.023**	0.0033**	-0.0013**	0.0049**
	(9E-5)	(0.0001)	(9E-5)	(0.0001)	(0.0005)	(0.0007)	(0.0005)	(0.0006)
θ_1	--	--	--	--	3.5E-6**	3.4E-6**	3.0E-6**	2.7E-6**
					(2E-7)	(2E-7)	(2E-7)	(2E-7)
θ_2	--	--	--	--	0.0001**	9E-5**	0.00011**	9.7E-5**
					(2E-5)	(2E-5)	(2E-5)	(2E-5)
θ_3	--	--	--	--	-4E-5**	-0.00022**	-3E-5**	-0.00022**
					(9E-6)	(2E-5)	(9E-6)	(2E-5)
γ [3]	124	118	123	114	126	134	125	129
θ	0.0037	0.0042	0.0038	0.0046	0.0035	0.0026	0.0036	0.0032
$R2$ [4]	0.41	0.46	0.45	0.53	0.43	0.50	0.46	0.56
F test	1643	--	402	--	1023	--	371	--
Hausman-Wu Test		2.95		6.19		15.60		14.16
DF for Hausman Test		2		17		5		20
(prob.>m)	--	(0.23)	--	(0.99)	--	(0.008)	--	(0.82)

** significant at the 1 percent level or better, *significant at the 5 percent level or better

(1) C= Monthly climate variable; SP= Percent of population of Spanish origin; AAP: 60-year average annual precipitation; M=Income per capita per day; P= Average price per gallon; M, SP, and AAP are time invariant.

(2) When dummies are included the intercept corresponds to the month of December, year 1985

(3) When dummies are included, all dummies are used to calculate γ

(4) Adjusted R-squares are reported for OLS regressions. For GLS, a generalized version is used [Buse (1973), SAS ETS (1996)].

Table 9: Stone-Geary: Nonlinear Regressions Results

$$Q = (1-\theta)(\alpha_o + \alpha_1 C + \alpha_2 SP + \alpha_3 AAP) + \theta \frac{I^*}{P} \text{ (1)}$$
$$\theta = \theta \text{ or } \theta = \theta_o + \theta_1 C + \theta_2 SP + \theta_3 AAP$$

Regression Type	OLS	GLS	OLS	GLS	OLS	GLS	OLS	GLS
Sample Size	9430	9430	9430	9430	9430	9430	9430	9430
$\theta=f(C, SP, AAP)$	No	No	No	No	Yes	Yes	Yes	Yes
Month and Year Dummies	No	No	Yes	Yes	No	No	Yes	Yes
α_0 (2)	39.87**	28.90**	59.89**	43.80**	109.08**	35.00**	117.44**	34.68*
(approx. st. error)	(3.37)	(11.53)	(4.79)	(12.95)	(7.36)	(14.08)	(8.07)	(15.07)
$\alpha 1$	0.076**	0.077**	0.070**	0.071**	0.034**	0.036**	0.035**	0.037**
	(0.0013)	(0.0011)	(0.0028)	(0.0024)	(0.0029)	(0.0023)	(0.0038)	(0.0031)
$\alpha 2$	-0.59**	-0.55	-0.54**	-0.46	-1.82**	-1.57**	-1.85**	-1.55**
	(0.098)	(0.48)	(0.095)	(0.53)	(0.22)	(0.52)	(0.21)	(0.56)
$\alpha 3$	-1.46**	-1.36**	-1.44**	-1.29**	-1.01**	1.32**	-1.02**	1.40**
	(0.058)	(0.29)	(0.057)	(0.32)	(0.13)	(0.35)	(0.13)	(0.37)
θ or θ_0	0.0037**	0.0042**	0.0038**	0.0046**	-0.0023**	0.0031**	-0.0013**	0.0048**
	(9.5E-5)	(0.00014)	(9.3E-5)	(0.00014)	(0.00054)	(0.00066)	(0.00053)	(0.00063)
θ_1					3.5E-6**	3.4E-6**	3.0E-6**	2.7E-6**
					(2.2E-7)	(1.8E-7)	(2.2E-7)	(1.7E-7)
θ_2					0.00010**	9.1E-5**	0.00011**	9.7E-5**
					(1.6E-5)	(1.8E-5)	(1.6E-05)	(1.7E-5)
θ_3					-3.5E-5**	-0.00021**	-3.2E-5**	-0.0002**
					(9.2E-6)	(1.6E-05)	(9.0E-6)	(1.5E-5)
γ (3)	124**	118**	123**	114**	126**	134**	125**	128**
	(3.45)	(11.00)	(5.16)	(12.57)	(7.09)	(13.54)	(8.09)	(14.61)
θ	0.0037**	0.0042**	0.0038**	0.0046**	0.0035**	0.0027**	0.0036**	0.0032**
	(9.4E-5)	(0.00014)	(9.3E-5)	(0.00014)	(0.00052)	(0.00053)	(0.00050)	(0.00046)

** significant at the 1 percent level or better, *significant at the 5 percent level or better

(1) C= monthly climate variable; SP= Percent of population of Spanish origin; AAP: 60-year average annual precipitation; M=Income per capita per day; P= Average price per gallon; M, SP, and AAP are time invariant.

(2) When dummies are included the intercept corresponds to the month of December, year 1985.

(3) When dummies are included, all dummies are used to calculate γ and its standard error.

CHAPTER 4

Numerical Simulations and Results

Dynamic Programming is the preferred solution method for discrete inter-temporal problems. In chapter 2, we developed the model and discussed the anticipated theoretical results. In this chapter, we find numerical solutions to the optimization problem; i.e., we approximate a solution function for g and s as a function of the current state of the system (H_t,M_t), assuming that next period's use of groundwater will be the same function of next period's state (H_{t+1},M_{t+1}). All functional forms for the model detailed in chapter 2 need to be explicitly specified and parameterized. We use data from the city of Houston, Texas, to parameterize the model. To facilitate the search for a numerical solution the problem is rescaled to per capita benefits and quantities in billion gallons . The benefit function is specified using the results of chapter 3 and Houston averages. The transition equation for water levels as a function of groundwater use is estimated using thirty years of well and groundwater use data. The transition equation for the minimum historic water levels (M), which determines subsidence, is approximated with a differentiable function of previous values of M and current water levels. Population growth is approximated using a linear trend to fit past census data since 1930 and projections from the City of Houston. The groundwater cost function is parameterized using an exact engineering relationship and an average price of energy. Surface water costs are parameterized using results from the literature. The distance function used in the surface water cost function is expressed implicitly using results from chapter 2 and Houston densities, city limits, and density gradients. The change in spatial parameters as population increases is approximated with continuous functions which are fit to historical data. Methodology issues concerning the numerical approximation of the

unknown value function and the optimization techniques are presented in section 2. The simulation is carried out using the functional forms and parameter values of section 1. We then perform sensitivity analysis to understand the impact of the important parameters on the solution.

1. Detailed Functional Forms and Parameterization Using the Case of Houston, Texas.

The next subsections detail how functional forms and parameters are obtained. A table is provided at the end of the section that summarizes the results (table 10, page. 88).

1.1 Transition Equations and Bounds of State Variables

1.1.1. Change in water levels

The transition equation for state variable H (hydraulic head or, equivalently, water levels), is of the form:

$$H_{t+1} = H_t + A - Bg_t$$

where A and B are constant parameters, and g is groundwater use in thousands of gallons. For data on water levels, we use the Texas Water Development Board (TWDB) database which includes records on all Texas wells and water level measurements. We only consider wells reported as belonging to the Evangeline aquifer, including those reported as drawing on the "Chicot-Evangeline" aquifer.[66] We used the Tiger Census map system to locate the wells according to their latitude and longitude. We find that the wells are scattered around the area. The wells used in the analysis are public supply wells and observation wells. We make sure that we only use wells for which water levels were static when the measurement was made (i.e., the well has not just been pumped, which would cause a temporary drop

[66] Even though our analysis will not be limited to this area, we choose wells in Houston and vicinity within the following limits: from South to North, between 29.37.00 and 29.51.50 of latitude and from East to West between –95.13.00 and -95.30.16 of longitude (latitude and longitude data are given in decimal numbers rather than degrees since this is the way they are recorded in the TWDB database). This corresponds to an area including the 610 Loop plus approximately four miles around it.

in water level due to the cone of depression created by pumping). Since the reporting of water levels in wells is very irregular, partly due to the fact that it needs to be measured after a period of no pumping, the data is very unbalanced. Between 1960 and 1995 we have an average of 70 wells in the sample with a minimum of 44 and a maximum of 89. An average of all the observations in a year is used for each well and then averaged out over all the wells. The reporting is heavily biased toward the end of the year so we consider the average to be an end-of-the year observation. For water use, we use data from the TWDB, reported by the City of Houston, which includes water production data broken down by source (groundwater and surface water). The data set gives groundwater use for the city of Houston from 1962 to 1995. Since Houston has expanded over time though annexations, an increase in water use for Houston does not directly imply an increase of pumpage for the area. To account for this problem, we look at groundwater use per connection and multiply this estimate by the number of connections at the beginning of the period (241,093 connections in 1962). In effect, we are looking at a fixed sample area, corresponding to the city of Houston in 1962, and we are estimating how the increase in usage in that area affects overall water levels, implicitly assuming that all other uses in and outside the area always increase proportionately.[67] The use of end-of-period number of connections would not change the results.

We run a simple regression of water levels on a constant and total groundwater production in the year. The following relationship is estimated:

$$H_{t+1} = H_t + \underset{(1.806)}{15.7} - \underset{(-2.074)}{0.424}\, g_t \qquad (4.1)$$

where g is water use per year in billion gallons and t-ratios are in parentheses.

[67] Later, in the Houston case study in chapter 5, we see that, at least, the population and water demand in the greater area including Harris and Galveston counties is expected to grow at the same rate as the Houston sample. See Figures 50 and 51 p. 123.

The F test on the regression is 4.3 (a p-value of 0.05). The R-square is 0.13. The signs are as expected and the coefficients are significant with significance levels better than 0.1 for the intercept and 0.05 for the slope coefficient. We suspect that the low R-square is due mostly to measurement errors.[68]. More research, beyond the scope of this dissertation, would have to be done to get a more reliable relationship between water levels and total groundwater withdrawals.[69]

The amount of water that can be withdrawn each year and leave the water levels unchanged is A/B. From the previous estimated equation this number is approximately 37 billion gallons per year (114,000 acre feet or 102 million gallons per day (mgd)). Any usage above the recharge causes water levels to drop. This estimate seems reasonable and it is useful as there is no reliable estimate on the safe yield of the Evangeline aquifer.[70]

1.1.2. Changes in the minimum historic head:

The transition equation for subsidence levels, as specified in chapter 3 is

$M_{t+1} = M_t$ when $H_{t+1} \geq M_t$;

$M_{t+1} = H_{t+1}$ otherwise.

This equation can be approximated using the following continuous function:

$$M_{t+1} = \frac{M_t}{\left((M_t/H_{t+1})^a + 1\right)^{1/a}},$$

The function can be rewritten in terms of the control by substituting $H_{t+1} = H_t + A - Bg_t$.

---

[68] Measurement errors is a common problem with this type of data. The problems arise mostly because the timing of measurements is different for each wells, which have to have not been pumped for a period of time allowing water levels to go up to their static level.

[69] Since we will be using the Stone-Geary demand parameters estimated from residential data mostly, one concern is how much of the water use reported by Houston is for residential use, and how much is for industrial use. The data indicate that in 1995, 65.9 percent of the water use reported as municipal to the TWDB was for residential use (including apartments), 31.4 percent was for commercial use, and 2.7 percent for industrial use. In 1996 the percentages were 65.4, 34.6, and 0 percent, respectively. Since residential usage is dominating, we can use this transition equation with the benefit function derived from last chapter results. We keep in mind, however, that the estimation is carried out with this restricted usage compared to the full usage in the area. For our results to be correct, we are implicitly requiring that other uses follow the same trend as municipal usage over time so that the proportion of other uses to municipal use stays constant.

[70] Note that the recharge comes from the outcrop about sixty miles northwest of Houston. Since in takes many years for the water to travel from the outcrop to the Houston area, we cannot directly relate recharge to rainfall. Effective recharge, however, would be lower in periods of drought when farmers use the groundwater to supplement surface sources. We do not consider this here and treat the recharge as fixed and exogenous.

When a is large, this function closely approximates the conditions above. It returns M_t if $H_{t+1} > M_t$ and H_{t+1} otherwise. With $a = 100$, the function is differentiable and virtually indistinguishable from its non-differentiable counterpart. For Houston, water levels are negative (below mean sea level). a negative ($a = -100$) so that we get M_t if $H_{t+1} < M_t$ and H_{t+1} otherwise. In order to calculate our first order conditions, the following transition equation is used:

$$M_{t+1} = M_t \left[(M/H_{t+1})^{-100} + 1 \right]^{1/100} \tag{4.2}$$

which requires that H and M be nonzero.

1.1.3.Bounds for H and M :

Upper and lower bounds for H and M impose implicit constraints on the use of groundwater. To find the upper bound, we use the earliest data points we have for water levels in the Evangeline aquifer, which show an approximate depth to water of 50 feet in 1930. As the average land surface elevation is about 55 feet, this is approximately 5 feet above mean sea level. We take $\widetilde{H} = -0.5$ (avoiding zero) to be the upper bound for H. For the lower bound, we have two choices. We can take the minimum water level before the aquifer is exhausted, or the water level at which the aquifer becomes a non-confined aquifer at atmospheric pressure. Because the Evangeline formation is a confined aquifer, wells are drilled deep enough to reach inside the aquifer but water levels in the wells will rise above the upper confining layer. When the water level drops below the level of the upper confining layer, the aquifer is being "dewatered", meaning that it is not under confined conditions. When the aquifer becomes unconfined, pressure levels remain constant but the available volume of water changes. The relationship between use and drawdown would be quite different for an unconfined aquifer. Moreover, dewatering the aquifer will cause significant geological changes to the formation (in terms of permeability, storage capacity, etc.). Also,

some wells are not drilled deep enough to take full advantage of the aquifer capacity and re-drilling wells would impose a large fixed cost that we do not take account of in this model. Therefore, we choose the lower bound for H to be the lowest level to which the aquifer could and would be pumped without causing dewatering. This corresponds to the level of the upper confining geological unit. For the Evangeline aquifer, this corresponds to about 500 to 600 feet below mean sea level in the Houston area.[71] From the TWDB wells database we have data on well depth for each well in the sample. The minimum well depth is 668 feet and the maximum well depth is 2580 feet. We place the minimum water level at 600 feet.

Since M never increases, we do not have to worry about upper bounds. The lower bound for M is the same as the lower bound for H, so $\underline{M} = \underline{H} = -600$.

1.1.4. Population growth and bounds:

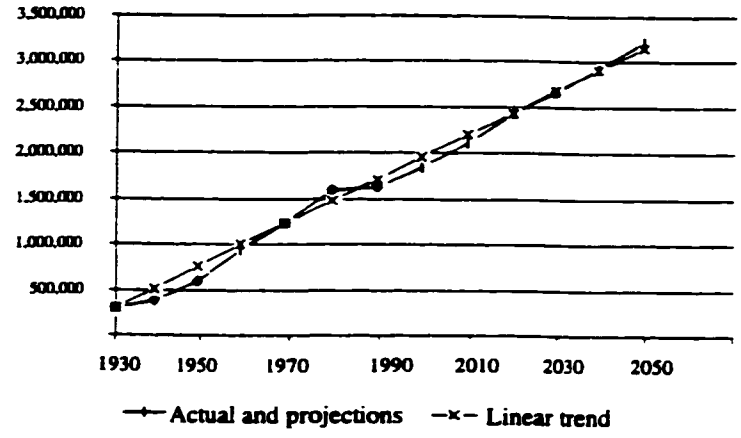
Our final state variable is population, N. We use population data for the city of Houston since 1930 from the Census and projections made by the TWDB. These projections give us a linear growth path from 1990 to 2050. The projected population for Houston in 2050 is 3.22 million. We choose 1930 as our starting date because water levels were rather stable at the time and close to their maximum. We find a growth function that approximates the actual and projected population data with a linear trend. We start with a population N_0 of 300,000 in 1930 , that grows arithmetically with $n = 0.08$ (+24,000 per year). In year 2050 it reaches 3.2 million. We choose the year 2050 for the start of the time invariant state when population reaches $N_{121} = \overline{N} = 3,180,000$, and stops growing. Figure 11 plots the proposed linear trend against the data (Census and TWDB estimates).

[71] Figure A1 in the Appendix shows a cross section of underground formations below Grimes, Montgomery, Harris, and Galveston counties. Houston is approximately in the middle of that cross section which shows the Evangeline aquifer between 600 feet and 2000 feet below mean sea level.

The growth path of population is given by:

$$N_t = N_0(1 + n(t-1)) = 300,000(1 + 0.08(t-1)) \tag{4.3}$$

Figure 11: Proposed Trend Against Actual and Projected Houston Population



1.2 Benefits: Gross Consumer Surplus

We use the Stone-Geary results from chapter 3. Recall that γ represent the subsistence parameter or "conditional water threshold" in gallons per capita per day, θ is used to represent the marginal budget share parameter, and I^* is per capita income per day.

Gross consumer surplus is defined as:

$W(w) = \int_0^w f(z)dz$ where $f(w)$ represents the continuously differentiable inverse demand for water. We check the regularity conditions for consumer surplus to be defined. In particular, we need $\infty > \lim_{w \to 0} \{-wf'(w)/f(w)\} > 0$.

Using the simplified Stone-Geary function from chapter 3, the inverse demand is $f(w) = \frac{\theta \cdot I^*}{w-(1-\theta)\gamma}$. We have $-wf'(w)/f(w) = \frac{w}{w-(1-\theta)\gamma}$. The condition is not satisfied since $\lim_{w \to 0} \frac{w}{w-(1-\theta)\gamma} = 0^-$. The correction is simple. We need to replace the zero limit with γ to obtain $\lim_{w \to \gamma} \frac{w}{w-(1-\theta)\gamma} = \frac{1}{\theta}$, which satisfies the condition for any level of γ. We constrain per capita water use to be at least equal to γ.

The gross consumer surplus for our Stone-Geary specification of water demand is

$$\int_0^w \frac{\theta \cdot I^*}{v-(1-\theta)\gamma} dv = k(\gamma) + \int_\gamma^w \frac{\theta \cdot I^*}{v-(1-\theta)\gamma} dv = k(\gamma) + \theta I^* \{ln[w-(1-\theta)\gamma] - ln(\theta\gamma)\}$$

where $k(\gamma)$ is a constant surplus up to $w = \gamma$. The value of k does not affect per capita marginal benefits and, therefore, will not affect the optimal policy solution. However, a positive value for k is assumed so that net benefits are positive. We cut the demand function at a price of $Pmax = \$7$ per thousand gallons (see chapter 3, figure 10) and $k(\gamma)$ is assumed to be $Pmax * \gamma$.[72] From chapter 3, we use θ and γ as functions of the exogenous variables: percent Spanish origin (SP), a climate variable (C), and average annual precipitations (AAP). We use 1990 Census data for SP, and average annual precipitation (1943-1995) from NOAA. We leave the C variable (number of days without significant rainfall $\times$ average temperature of the month) at the 1980-85 Texas average.[73] The values used are: $SP = 22.9$; $C = 1790$; $AAP = 43.7$. We assume that these variables do not change over time.

The coefficients estimated in chapter 3, using the non-linear GLS method with month and year dummies, and evaluated at the Houston means, are

$\gamma = 112.84$ gallons per capita per day (gcd)

$\theta = 0.0047$.

The average per capita income for Houston in the 1990 census is approximately $\$15,202$. Since most of our reference cost estimates are from 1975, we deflate this income with a consumer price index to reflect 1975 prices. This translates into a real income of approximately $\$6,250$ in 1975 dollars.[74] Since the time unit is a year, income does not need to be transformed, so $I^* = \$6,250$. θ is unit free. However, γ need to be in billion gallons per capita per year, which gives $\gamma = 112.84 \cdot 365/10^9 = 4.1187 \cdot 10^{-5}$. We want to see the impact

[72] The estimation of the threshold γ was conditional on the sample. We did not have prices above \$7 per 1000 gallons in the sample. As mentioned in chapter 3, with a higher price and given that adjustments in consumption of complementary goods can be made, we expect γ to be reduced.

[73] The C variable in our demand estimation was mostly important to capture seasonality. Since we only have annual data here, it's significance is reduced.

[74] CPI from the Bureau of Labor Statistics. 1975 index = 53.8, 1990 index = 130.7.

of a lower threshold γ, to understand the implications of possible non-price based demand management conservation policies, or greater reliance on demand related pricing. Real per capita income, assumed constant in the model, will continue to increase over the planning horizon and this trend is anticipated to increase water demand.

The gross benefit function per capita per year for quantities in billion gallons can be fully parameterized as follows:

$$W(w) = Pmax \cdot \gamma + \theta I^{*} \left\{ ln \left(\frac{s+g}{N} - (1-\theta)\gamma \right) - ln\left(\theta\gamma\right) \right\}, \qquad (4.4)$$

With $Pmax = 7 \cdot 10^{-6}$, $\theta = 0.0047$, $\gamma = 4.1187 \cdot 10^{-5}$, and $I^{*} = 6,250$.

1.3 Groundwater Costs

We ignore the fixed costs associated with well investment and maintenance. These costs do not increase with depth and would just add a constant term that depends on the number of wells. Since we do not let the number of wells change (no investment) and maintenance costs are negligible compared to pumping costs and distribution costs, we ignore these costs in the analysis. Distribution costs common to groundwater and surface water do not affect costs differentials and therefore, only affect the quantity of water use, not the choice of source. In the second simulation, we will add a linear cost function of both groundwater use and surface water use. This will only reduce overall water use and therefore can be considered as a reduction of benefits rather than an increase in the specific costs of groundwater or surface water. For the first simulation, we focus only on the energy cost of pumping and subsidence costs.

1.3.1.Groundwater Pumping Costs

As indicated in chapter 2, pumping costs increase linearly with depth according to the following relationship

$C_p(g, H) = g[c_0 - c_1 H]$.

This relationship is driven by the amount of horsepower needed to lift Q gallons of water per minute from a vertical distance of $(S_L - H)$ feet (where S_L is the level of the land surface), given a pump and motor efficiency Ep:[75]

$HP = \frac{Q(S_L - H)}{3960 Ep}$

We take Q to be the rate of pumping of large municipal water wells which is about 2000 gallons per minute. Taking an average pump efficiency of .85, motor efficiency of .9, Ep averages around .75. We can translate horsepower into energy consumption in kilowatts

$KW = 0.7454 \frac{2000(S_L - H)}{2970} \approx 0.502\,(S_L - H)$

Multiplying by the price of energy (Pe) per KWH and multiplying by the number of hours needed to pump g gallons of water at a rate of 2000 gallons per minute[76]

$C_p(g, H) = Pe \cdot 0.502\,(S_L - H)\,\frac{g}{2000}\frac{1}{60} \approx Pe \cdot 4.18 \cdot 10^{-6} g\,(S_L - H)$

given that S_L, the elevation of land surface in feet from mean sea level (MSL) is on average 55 feet in the Houston sample, and transforming g to be in billion gallons, we get

$$C_p(g, H) \approx g \cdot Pe\,(230 - 4.18H) \cdot 10^{-6} \tag{4.5}$$

We do not have a function for surface water costs that includes the price of energy explicitly and so choose a constant price of energy.[77] For quantities in billion gallons and a price of energy of 5 cents per KWH, the pumping cost function is:

$$C_p(g, H) \approx g \cdot Pe\,(11500 - 209H) \tag{4.6}$$

This equation yields a cost that varies between 5 and 7 cents per 1000 gal. when water levels go from -200 to -285, which is consistent with reported groundwater costs and water

---

[75] I wish to thank Tom Edgar, professor of Civil Engineering at the University of Wyoming for his input on determining the groundwater pumping cost function.

[76] Since the water comes from many wells, the number of hours could substantially exceed the number of hours in a year. However, we do have to assume that all wells are similar and pumped at the 2000 gallons per minute rate.

[77] We will see that with subsidence costs, the impact of energy prices on the solution becomes negligible.

level declines between 1970 and 1980.[78] Given the geology of the Evangeline aquifer, the impact of changing water levels is a minor influence on changing groundwater costs. For the depths allowed, groundwater pumping costs vary between 2.2 to 13.7 cents per 1000 gal., the largest cost still below the cost of treated surface water.[79] The pumping externality alone would not justify using surface water until geological constraints are binding. If the price of energy is increased, the pumping externality will become more significant.[80] This shows the importance of subsidence costs as, in there absence, and considering only cots differentials, surface water would not be used until the aquifer is depleted.

1.3.2. Groundwater Subsidence Costs

We could find no functional estimates of subsidence costs. Few studies have looked at the actual cost of subsidence, which is difficult to estimate. A study by Warren et al. (1975) estimated subsidence related losses in the Houston-Baytown area at $113.6 million between 1943 and 1973. Between 1969 and 1973, losses were estimated at $73 million, an average of $14.6 million per year. This leaves $26 million for the period 1943-1968, an average of $1.04 million per year.[81] This data is consistent with our increasing costs assumption. We find that a quadratic function

$$C_s(M) = \bar{c}\,(H_0 - M)^2 \tag{4.7}$$

---

[78] Most of our average costs estimates come from Warren et al. (1975). We also got some information in Teutsch (1978). Although it is difficult to tell the reference year in these sources, we assume all prices to be in 1975 prices.

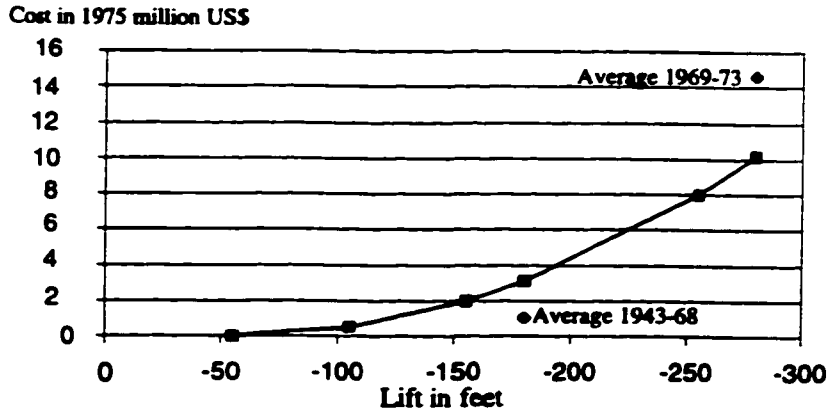
[79] This is consistent with conversations we had with water professionals in Houston, who do not consider that the change in pumping costs due to change in depth is of significant concern, unless the professional in question has been around for a very long time. For instance, Robert Gabrysch, retired but still active at the Houston Galveston Subsidence District, formerly with the Texas branch of the US Geological Survey, believed that depth did make a significant difference in pumping costs. However, Gary Binderim (ST Laboratories, a Houston engineering firm dealing with the production of groundwater for Municipal Utility Districts) thought that depth did not matter, unless screens had to be lowered or wells redrilled (interviews in 1997-1998).

[80] Even with a doubling of the price of energy, the cost of groundwater would only reach $.27/1000 gal. at -600 feet. This would still not be sufficient to justify the switch to surface water, especially when we note that surface water also uses energy. In places where water levels can drop to much deeper levels, such as in El Paso, TX, where the supply of groundwater comes from deep aquifers, the pumping externality might be enough to justify surface water use, and it has.

[81] Estimated costs come from data from public agencies and surveys to private individuals and firms. Costs were assumed to be reported in current dollars.

is a good candidate for the subsidence cost function. Subsidence costs increase at an increasing rate for the full range of possible head values and are zero when the aquifer has not been mined. Fitting 4.7 to the median water levels for the two periods reported by Warren et al., 1975 (125 ft. for 1943-68 and 225 ft. for 1969-73, which correspond to a lift of 180 ft. and 280 ft. respectively, given an average land surface elevation of 55 ft.) results in costs of about \$3 million and \$10 millions per year respectively when $\bar{c} = 200$. In the simulation the level of zero subsidence is taken to be at H_0, the upper bound of H. With $\bar{c} = 200$, given minimum water levels at $M = -263$, and given groundwater use at 55 billion gallons, the average cost of subsidence per thousand gallons of groundwater in 1975 is about 25 cents/1000 gallons. This is above the average cost of surface water at the time (around 16 cents per thousand gallons). Figure 12 shows the graph of our subsidence cost function and the points corresponding to observed data. These estimates of subsidence are based on one study and we are uncertain about their accuracy. Since the $\bar{c}$ coefficient is expected to have an important impact on the solution, we will perform sensitivity analyses on this parameter.

Figure 12: Subsidence Cost Function and Observed data



1.4 Surface Water Costs

Surface water costs include, mainly, acquisition and treatment costs, transmission costs (from the treatment plant to the distribution system), and distribution costs. Again, distribution costs are common to surface and groundwater provision and therefore do not affect costs differentials. We only consider linear distribution costs in a "reduced benefits" case in a second set of simulations. We focus here on treatment costs and transmission costs. We separate surface water costs into two functions: a treatment cost function f_1, and a transmission costs function f_2. Total surface water costs are $f_1 + f_2$. The transmission costs function was developed by Clark and Stevie (1981).[82] The acquisition and treatment components of the function are also in Clarke and Stevie but they are common in the literature. We add a penalty function for quantities of water beyond maximum capacities as explained below.

1.4.1. Surface Water Treatment Costs

The treatment cost function is of the form:

$$\text{Treatment costs} : f_1 = D_1 (s)^{\alpha_1} + e^{s-k} \tag{4.8}$$

where e is the base of natural logarithm. The scale factor for treatment plants is estimated at $\alpha_1 = 0.77$ (Clarke and Stevie, 1981; DeLucia et al., 1978). The constant D_1 was estimated at 513 for quantities in million gallons (1977 estimate). For s in units of billion gallons, D_1 needs to be rescaled by a factor of 1000^{α_1}, this translates into $D_1 = 104,741$. However, this coefficient implies treatment costs of 4 cents per 1000 gal., which would grossly underestimate the cost of treatment according to several sources. The cost of conventional treatment was estimated to be 13.7 cents per thousand gallons for a 100 mgd plant in 1980.[83]

---

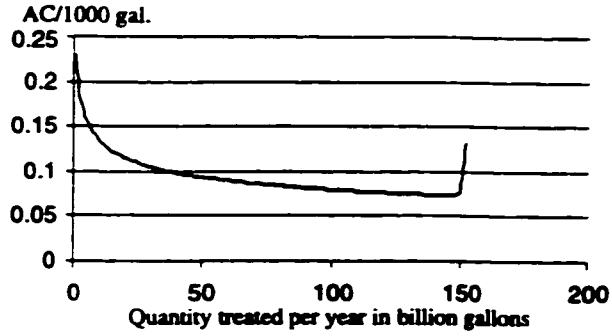
[82] The transmission costs in Clarke and Stevie include distribution costs which should also be included in groundwater costs. It would be better to take distribution costs out of the surface water cost function so the difference in groundwater and surface water cost is not overstated. We need to keep this in mind in the interpretation of our results.

[83] Robert Clark , American Society of Civil Engineers, "Cost estimating for Conventional Water Treatment" Journal of the Environmental Engineering Division, Proceeding of the American Society of Civil Engineers, Vol. 108, No EES, October 1982.

This is approximately equivalent to 9 cents per1000 gal. in1975 prices.[84] Surface water use data for 1975 (approximately 55.4 billion gallons of surface water), $D_1 = 229,280$ imply a treatment cost of approximately 9.1 cents/1000 gal. in 1975.[85]

As treatment facilities are limited to one location in this model, we put a bar on the quantity of surface water that can be produced with increasing returns. We impose a "penalty" function, which limits the range of increasing returns to the observed maximum capacity of a single location water treatment facility. The data indicates that for any location in Houston, there is a maximum capacity of about 400-500 mgd (see Appendix, figure A2). In other words, capacity can be increased at any location until it reaches approximately 400-500 mgd (400 mgd is equivalent to 146 billion gallons per year). Any surface water use beyond this capacity would increase average costs sharply. This is, in effect, imposing a capacity constraint on surface water, but has the important advantage of avoiding corner solutions. The parameter k is chosen so that the average cost function for surface water is minimized around $x = 146$ billion gallons and increases sharply for quantities greater than 146 billion gallons per year. With $k = 136$, the average cost function reaches a minimum at $x = 145.9$ (figure 13) This will prove important when we solve for the control values at the end period.

Figure 13: Average Treatment Costs



[84] We use the The Bureau of Labor Statistics CPI. The CPI values 82.4 in 1980 and 53.8 in 1975.

[85] In the following, all calculated average costs use the data on surface water use in 1975.

1.4.2. Surface Water Transmission Costs

The transmission costs to a point in the service area are of the form:

$D_2 Q_p^{\alpha_2}$, with $\alpha_2 < 1$

where Q_p is the water distributed at a point, which is equal to per capita consumption times population density at that point. We assume that per capita consumption is the same across the city whether residents are provided with groundwater or surface water, and use a per capita consumption of $\frac{s+g}{N}$. The transmission cost function is integrated over the area of surface water service with limit of integration d determined by the proportion of surface water to total water use and the spatial characteristics of the city, as derived in chapter 2. An exponentially decreasing density is assumed, but the city need not be circular. We use $\overline{\Theta}$ to indicate the limit of integration for the geometry of the city. We can now express the cost function in terms of spatial variables K (density at the center), λ (density gradient), and d (maximum distance surface water needs to be transported). The transmission cost function is integrated over the area where surface water is used. The following function for transmission costs is derived by Clark and Stevie (1981):

$$\text{Transmission costs: } f_2 = \frac{\overline{\Theta} D_2 \left(\frac{s+g}{N} K\right)^{\alpha_2}}{(\alpha_2 \lambda)^2} \left[1 - e^{-\alpha_2 \lambda d} (1 + \alpha_2 \lambda d)\right] \qquad (4.9)$$

We use the analytical results of chapter 2 to specify d as the maximum distance to which surface water needs to be transported. In chapter 2, we derived the minimum distance d from the center such that a proportion x of the population is included in a radius of d. Given an exponentially decreasing density function for a circular city, population within a radius of d can be expressed as

$$K \int_0^{\overline{\Theta}} \int_0^{d} e^{-\lambda s} s \, ds \, d\theta = \frac{\overline{\Theta} K}{\lambda^2} \left[1 - e^{-\lambda d} (1 + \lambda d)\right]. \qquad (4.10)$$

From equation 4.10, the total population in a city of radius l is

$$\frac{\overline{\Theta}K}{\lambda^2}\left[1-e^{-\lambda l}(1+\lambda l)\right] \tag{4.11}$$

Therefore, to reach a proportion x of the population, we need d such that

$$1-e^{-\lambda d}(1+\lambda d) = x\left[1-e^{-\lambda l}(1+\lambda l)\right] \tag{4.12}$$

This function cannot be solved explicitly for d, but, fortunately, the form of the surface water cost function allows us to leave the solution for d (given x) in its implicit form given in equation 4.12.

If condition 4.12 is true for any λ, it must be true for $\alpha_2\lambda$, since α_2 is a constant. We substitute condition 4.12 into equation 4.9, and replace x by $\frac{s}{s+g}$, given our theoretical results of chapter 2. Together, these transformations give us

$$f_2 = \frac{\overline{\Theta}D_2 s\,(s+g)^{\alpha_2-1}}{(\alpha_2\lambda)^2}\left(\frac{K}{N}\right)^{\alpha_2}\left[1-e^{-\alpha_2\lambda l}(1+\alpha_2\lambda l)\right] \tag{4.13}$$

Clarke and Stevie (1981) estimated $\alpha_2 = 0.82$.[86] The constant D_2 was estimated at 408 (1977 estimate for quantities in million gallons). D_2 needs to be rescaled by a factor 1000^{α_2} to accommodate quantities in billion gallons. The resulting $D_2 = 117,668$, implies transmission/distribution costs of 17 cents/1000 gal in 1975. Since this estimate includes distribution costs, which we have ignored in groundwater cost estimates (or which we will consider as linear costs common to both sources) we reduce this coefficient so that the average costs treatment and transmission of surface corresponds to a 1976 estimate of 20.3 cents/1000 gal. for average cost of treated surface water delivered at a major distribution line (Teutsch ,1977). We choose $D_2 = 83,176$, which gives us an average cost of transmission

[86] Clarke and Stevie's estimate for α_2 was the only one available. Note that, in their paper $\beta = 82$ (their β is our α_2) is, in fact, an obvious typographical error; we reproduced their cost simulations and found that the more credible $\beta = 0.82$ was used.

of 12 cents/1000 gal. for quantities of surface water in 1975, which, added to the average treatment costs yields an average cost of treatment and transmission of 21 cents/1000 gal. in 1975.

As expected, these parameters produce increasing returns in treatment costs ($\alpha_1 = 0.77$) and transmission costs ($\alpha_2 = 0.82$), but decreasing returns to dispersion (as distance to service increases, holding quantity of water constant, average transmission costs increase). Specifically, unit transmission costs are decreasing in s, decreasing in g, increasing in l, increasing in K, and decreasing in λ. If service area does not increase then increasing returns dominate. An increase in service area soon makes unit surface water costs increase.

1.5 Spatial Parameters

Another issue relates to the spatial nature of the model and the full parameterization of the surface water cost function is how the parameters λ, l, and K change as N changes. To determine the impact of a change in N, it is important to see how these variables change as population increases. Mieszkowski and Smith (1991) provide some estimates of λ and values of K and l for the City of Houston between 1950 to 1980. Within the period, density at the center, K, went from about 15,000 per square mile to 10,500 per square mile. The metropolitan area density gradient, λ, went from 0.283 to 0.148. The distance to city limits changed from 14 miles to 30 miles. It is common to see population densities to decline at the center with density gradient declining and urban areas spreading further out.

There is an exact relationship between K, λ, l, and N which needs to be maintained in our simulations. Population N is already determined each period by the arithmetic growth path given earlier. Every period, we need

$$\frac{\overline{\Theta}K}{\lambda^2}\left[1 - e^{-\lambda l}(1 + \lambda l)\right] = N \tag{4.14}$$

From this relationship, the estimates in Mieszkowski and Smith (1991), and the actual

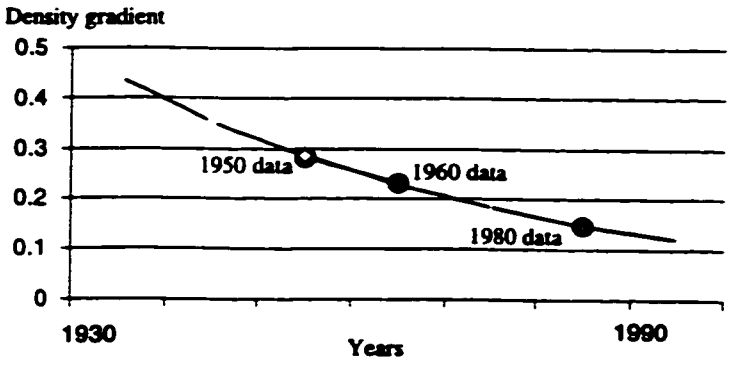
population in the City of Houston in 1950 and 1980, we deduce that a circular city ($\overline{\Theta} = 2\pi$) would over-estimate the Houston population. We choose $\overline{\Theta} = 1.12\pi$, a little greater than a half circle, to approximate the shape of Houston. With $\overline{\Theta} = 1.12\pi$, the numbers for K, and l, together with the estimated λ for 1950 and 1980 would give a population of 596,790 in 1950 (actual 596,163) and of 1.58 millions in 1980 (actual 1.59 million). For simplicity, we assume $\overline{\Theta}$ to be constant.[87]

The behavior of λ over time is extrapolated given past data. We assumed that λ decreases at a constant rate, which gives us the following growth equation for λ:

$$\lambda_{t+1} = \frac{\lambda_t}{(1+r_1)} = \frac{\lambda_0}{(1+r_1)^t} \tag{4.15}$$

As figure 14 below shows, if we start at $\lambda_0 = 0.44$ in 1930, and assume $r_1 = 0.022$, we get a nearly perfect fit with this function to the values calculated in Mieszkowski and Smith for 1950, 1960, and 1980.

Figure 14: Transition Equation for Density Gradient versus Observed Data



The behavior of l is also modelled using past data. Houston still has lots of space to

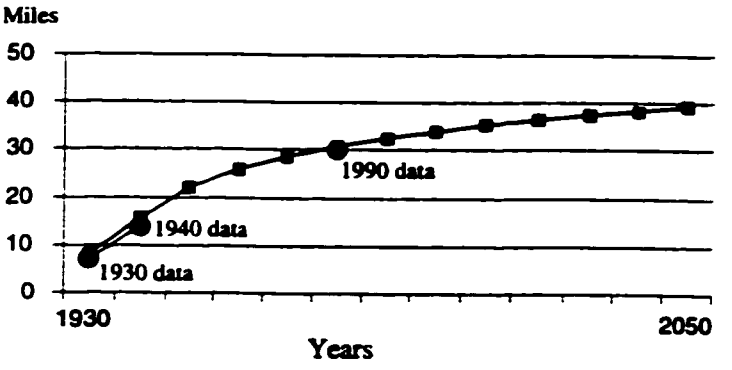
[87] $\overline{\Theta}$ was calculated as a parameter to fit our population data and spatial characteristics, without prior on its realistic value. It is possible that changing the value of $\overline{\Theta}$ serves as a correction to the choice of exponentially declining density function rather than some other density function. However, since the exponential density is accepted in the literature, and since it is the only one for which we have reliable estimates, we choose to not address the issue further.

grow in size, although it is difficult to anticipate how far it will have grown by 2050. Annexations have been the way Houston has typically grown so far, and more annexations are anticipated in the future. We assume that the spatial expansion of the city after 1940 follows the following relationship

$$l_t = r_2 l_0 ln(t/2) \tag{4.16}$$

In order to keep l increasing for low values of t, we assume a functional $l_t = l_0(t/1.3)$ for $t < 10$, which smoothly extend our general equation. With $l_0 = 8$ miles, and $r_2 = 1.2$, this function is consistent with past data and implies that city limits will be close to 40 miles in 2050 (figure 15).

Figure 15: Transition Equation for City Limits versus Observed Data



This leaves us with one variable K, which must be determined endogenously given the relationship established in 4.14. Every period we must have

$$K = \frac{N\lambda^2}{\overline{\Theta}\left[1 - e^{-\lambda l}(1 + \lambda l)\right]} \tag{4.17}$$

Given our predicted population growth, change in gradient and change in city limits, we would have a central density at the center of approximately 2550 inhabitants per square mile

in 2050.[88]

This condition can easily be substituted into the surface water transmission cost function to get

$$f_2 = s\frac{\overline{\Theta}^{1-\alpha_2}(0.408)(\lambda^2(s+g))^{\alpha_2-1}}{\alpha_2{}^2[1-e^{-\lambda l}(1+\lambda l)]^{\alpha_2}}\left[1-e^{-\alpha_2\lambda l}(1+\alpha_2\lambda l)\right] \tag{4.18}$$

1.6 Discount Factor and Summary

One last parameter needs to be specified, the discount factor β. There is always controversy about the choice of an appropriate discount factor. We choose to start the simulation with $\beta = .943$ which corresponds to a discount rate of approximately 6 percent per year. We will vary the discount rate in our sensitivity analysis section. Following is a summary table of all parameter values, functional forms, initial values, and bounds used in the first simulation run.

2. Simulation

2.1 Methodology

In order to avoid large values that would create problems in the numerical algorithm, we restate our problem in per capita terms. The dynamic program is written as:

$$V(H,M,\overline{N}) = \max_{g,s}\left\{W - \frac{C_p - C_s - f_1 - f_2}{\overline{N}} + \beta V(H',M',\overline{N})\right\} \tag{4.19}$$

Where $H' = H + A - Bg$,and $M' = min(M,H')$.

The maximization needs to be carried out subject to several constraints:

First, in order to keep H at or above its lower bound, we must have at all times

[88] It seems doubtful that central densities would continue to decline at the same rate in the next fifty years. However, densities are projected to be more uniform and city limits to expand. Given our density function, a lower density at the center is necessary to obtain such patterns over time. This may be considered as accommodating for vacant land and population clusters.

Table 10: Summary of Parameters, Functional Forms, Bounds, and Starting Values

Parameter	Value	Description
A	15.7	Increase in hydraulic head (H) due to recharge
B	0.424	Marginal effect on H of pumping 1 billion gallons
n	0.08	Population increase coefficient (linear)
γ	4.1187e-5	Water use threshold in billion gallons per year
I^*	6,250	Average income per capita in 1975 dollars per year
$\overline{c}$	200	Subsidence costs coefficient
α_1	0.77	Treatment costs coefficient
α_2	0.82	Transmission costs coefficient
r_1	0.022	Rate of decrease in density gradient
r_2	1.2	Scaling coefficient for the increase in city limits
β	0.943	Discount factor corresponding to 6 percent discount per year

(Equation #) Type	Functional Forms, Bounds, Initial Values
(4.1) Water Levels	$H_{t+1} = H_t + A - Bg_t$
(4.2) Subsidence	$M_{t+1} = min\,(M, H') \simeq M_t \left[\left(\frac{M_t}{H_{t+1}}\right)^{-100} + 1\right]^{1/100}$
(4.3) Population Growth	$N_t = N_0\,(1 + n(t-1))$
(4.4) Benefits	$W = (7e+6)\gamma + \theta I^* \left\{ln\left(\frac{s+g}{N} - (1-\theta)\gamma\right) - ln(\theta\gamma)\right\}; \theta = 0.0047$
(4.6) Pumping Costs	$C_p = g\,(11500 - 209H)$
(4.7) Subsidence Costs	$C_s = \overline{c}\,(H_0 - M)^2$
(4.8, 4.18) Surface Water Costs	$f_1 = 229{,}280s^{\alpha_1} + e^{(s-136)}$, $f_2 = 83{,}176\frac{s\left(\lambda^2(s+g)\right)^{\alpha_2-1}\left[1-e^{-\alpha_2\lambda l}(1+\alpha_2\lambda l)\right]}{(1.12\pi)^{\alpha_2-1}\alpha_2^2\left[1-e^{-\lambda l}(1+\lambda l)\right]^{\alpha_2}}$
(4.15) Density Gradient	$\lambda_t = \lambda_0\,(1+r_1)^{-t}$
(4.16) City Limits	$l_t = r_2 l_0\,(t/2)$ for $t \geq 10$ and $l_t = l_0(t/1.3)$ for $t < 10$
Bounds on State Variables	$\underline{H} = \underline{M} = -600$, $\widetilde{H} = -0.5$ (feet from MSL); $\overline{N} = 3{,}180{,}000$
Initial Values (t=0 in 1930)	$H_0 = M_0 = -0.5$; $N_0 = 300{,}000$; $\lambda_0 = 0.44$; $l_0 = 8$

$$H + A - Bg \geq \underline{H} \tag{4.20}$$

Even though water levels cannot exceed $\widetilde{H}$, the upper bound on H does not impose a direct constraint on g. We only need to constrain water levels to not exceed $\widetilde{H}$ in the transition equation. This eliminates cost savings from reducing g below recharge if water levels are at their maximum. However, g need not be constrained, in case it is still optimal to reduce g beyond the point for which $H + A - Bg > \widetilde{H}$.

We need one additional constraint because the specific form of the benefit function assumes a minimum consumption of water (γ). Consumer surplus would not exist if the quantity of water consumed were below this threshold. The following constraint is added:

$$s + g \geq \gamma. \tag{4.21}$$

It turns out that this constraint is necessary for programming purposes (in the search for a solution, consumption levels below γ produce imaginary numbers, which then prevent from proceeding to find a solution). The constraint ??, however, will prove non-binding at all times and for all state levels.

We need the optimal controls g and s as a function of the current state of the system (H_t, M_t), assuming that next period's use of groundwater will be the same function of next period's state (H_{t+1}, M_{t+1}). In order to find such a function $G(H, M)$, we need to solve the first order condition for g, which includes the derivative of the unknown value function. We first specify a set of values for H and M for which the function $G(H, M)$ will be evaluated. Following Hartley (1996) and Gao, Hartley, and Sickles (1999), we use numerical techniques to solve the value function for a last (time invariant) period. Given that our problem involves two state variables, a tensor spline approximation is used. It requires

us to specify, for each state variable, a set of knots and collocation points at which a series of piecewise polynomials are fitted. Degrees of smoothness across these knots are also specified. We choose a spline of order 4 (cubic spline). Depending whether we anticipate a continuous or discontinuous first and second derivative of the value function, we impose different smoothness conditions. Continuity of the function, first derivative, and second derivative allows us to impose three smoothness conditions (single interior knots). When we expect discontinuity in the first derivative, we only impose one smoothness condition, which requires triple interior knots. The spline coefficients are estimated to minimize the sum of squared errors.[89] Our problem is defined over a triangular grid because M cannot be greater than H. We rescale one of the states to be a ratio of the other, so one of the state variables is H/M and the other is M. This allows us to use algorithms that require a rectangular grid.[90] For the optimization algorithm, we supply analytical gradients, which correspond to our first order conditions, which require the numerical derivatives of the approximated value function, and first derivatives of the constraints.

Before we can be sure that the numerical approximation will give the desired results, we need to make sure that the time invariant value function exists, is unique, and is well behaved so that the optimal solutions to the control problem also exist and are unique. The dynamic programming methodology, using Bellman's principle of optimality, has been shown to yield unique solutions comparable to the optimal solutions of the structural problem (objective

---

[89] The Spline toolbox in MatLab approximates a function to fit some known data points at a number of grid points (collocation points) given a smaller number of knots. The collocation points are calculated using Gauss points from an even grid. We choose four collocation points between each knot and make sure that the knots span an interval slightly greater than the collocation points to accommodate numerical approximation of next period values. The optimization toolbox handles the constrained optimization. See Hartley (1996), section 5, for a detailed explanation of a similar procedure with a single state variable.

[90] We tried to use H and M as states and fill the grid, first with a constant value function, then with linear interpolation. Even though the problem never uses the part of the grid where $M > H$, the approximation algorithm for the value function does. Solutions were extremely sensitive to the method we chose and mostly inconsistent. I must thank Peter Hartley for thinking through this problem and coming up with the scaling solution.

function given in chapter 2) under certain conditions.[91] A solution to the value function in the time invariant state is a fixed point of the mapping $T : T \cdot V = V$, where V is the value function and T is an operator on the space of bounded and continuous functions (T needs to satisfy Blackwell's sufficient conditions for a contraction mapping). Without the existence of this fixed point, the problem may not have a solution or the solution may not be unique. Sufficient conditions for existence and uniqueness require concavity of the objective function, convexity of the constraint set, and a discount rate less than one. If the objective function is not strictly concave, which is the case in our problem when increasing returns are dominating, then local optima may be obtained and more care needs to be taken in the search for the solution. To be more specific, we shall introduce some new notation.

Let σ denote the vector of policy variables (s, g) and let x be the set of all possible values for the state variables, $x = (H, M)$. Let $\sigma \in \Sigma$, where Σ represents the set of all possible values of σ, and let $x \in X$, where X represents the set of all possible values of x. let $\Gamma : \Sigma \times X \rightarrow X$ be a correspondence describing the feasibility constraints, so for each $x \in X$, $\Gamma(\sigma, x)$ is the set of possible values for the state variables next period if the current state is x and the policy variables are σ. Let B denote the objective function (gross benefits minus costs). B is a function mapping from $\Sigma \times X$ into R, where R is the set of real numbers. Finally $C(X)$ is the set of bounded and continuous functions with the sup norm: $\|f\| = sup_{x \in X} |f(x)|$. Using the mapping T defined above the dynamic programming functional equation 4.19 can be expressed as follows:

$(T \cdot V)(x_t) = \underset{\sigma \in \Gamma(x_t)}{max} \{B(\sigma_t, x_t) + \beta V(x_{t+1}(\sigma_t))\}$ which can implicitly be written in terms of the current and next period state variables (we use y to denote the next period state):

$(T \cdot V)(x) = \underset{y \in \Gamma(x)}{max} \{B(y, x) + \beta V(y)\}$,where it is understood that, though the transi-

[91] General proofs of the equivalence of the two approaches are given in Stokey and Lucas (1989). In the following we closely follow this work pp.77-87, directly referring to their theorems and verifying that their assumptions are satisfied in our problem.

tion equations for H and M, y is implicitly a function of x and σ, satisfying the constraint set so $y \in \Gamma(x)$.

By theorem 4.6 [Stockey and Lucas (1989), later referenced as SL, p.79], T has a unique fixed point V in $C(X)$ and for all V_0 in $C(X)$, we have

$\|T^n V_0 - V\| \leq \beta^n \|V_0 - V\|$, for $n = 0, 1, 2, ...$

Moreover, given V, the optimal policy correspondence

$G(x) = \{y \in \Gamma(x) : V(x) = B(x, y) + \beta V(y)\}$

is compact valued and upper hemi-continuous.

The proof requires that

(A1) X is a convex subset of R^l and Γ is non-empty, compact valued, and continuous (Assumption 4.3 in SL), and that

(A2) the objective function B is bounded and continuous and $0 < \beta < 1$ (Assumption 4.4 in SL).

We need to show that these two assumptions hold for our problem.

The first part of (A1) is straightforward, the state space is well defined with continuous real values over the range of possible water levels between $\underline{H}$ and $\widetilde{H}$. The second part require first, that there exists at least one possible value of y, the next period states, for any x in X, which is ensured by the fact that the transition equations are continuous and defined over the whole interval $[\underline{H}\ ,\ \widetilde{H}]$ and $[\underline{H}\ ,\ H]$ respectively, the latter consisting of at least of one point, $\underline{H}$. Also, the amount of groundwater that can be pumped is bounded below by the non-negativity constraint and above by 4.20, which is a bounded and continuously differentiable function. By the Implicit Function Theorem, the correspondence Γ is non-empty, compact valued, and continuous when X is a bounded and closed set (compact). Finally $B(x, \sigma)$ is a sum of continuous functions, and therefore continuous, and since the transition equations are continuous, the implied $B(x, y)$ is continuous . Given (A1) and the

compactness of X, B is also bounded so (A2) is satisfied.

Theorem 4.3 in SL ensures that if the above holds, the unique solution V satisfying 4.19 is the supremum function for the associated sequence problem. There exists at least one optimal plan and any plan generated by the correspondence $G(x)$ is optimal.

We add a corollary to further characterize the value function and optimal solution set. By theorem 4.7 in SL, if V is the unique solution to the dynamic programming model 4.19 in the stationary state, then V in strictly increasing in both state variables H and M. This requires two new conditions to be satisfied:

(A3) The objective function B is strictly increasing in current H and M (Assumption 4.5 in SL).

(A4) Γ is monotone in $x = (H,M)$ in the sense that $x \leq x'$ implies $\Gamma(x) \subseteq \Gamma(x')$ (Assumption 4.6 in SL).

It is easy to check that the net benefit function is increasing in both H and M since pumping costs are increasing in H and subsidence costs are increasing in M. Finally, $\Gamma(x) \subseteq \Gamma(x')$ for $x \leq x'$ since the higher the water levels the less constraining 4.20 is on groundwater use. Consequently, V is strictly increasing in both state variables.

Finally V is strictly concave and G is single valued and continuous if two final conditions hold (Theorem 4.8 and Assumption 4.7 and 4.8 in SL):

(A5) B is strictly concave, i.e. $B\left(\alpha(x,y)+(1-\alpha)(x',y')\right) \geq \alpha B\left(x,y\right)+(1-\alpha)B\left(x',y'\right)$, for all (x,y) and (x',y') in the graph of Γ, all $\alpha \in (0,1)$, and the inequality is strict if $x \neq x'$.

(A6) Γ is convex, i.e. $y \subseteq \Gamma(x)$ and $y' \subseteq \Gamma(x')$ implies $\alpha y+(1-\alpha)y' \in \Gamma\left(\alpha x+(1-\alpha)x'\right)$.

We note that an upper bound for surface water is not imposed explicitly but is ensured through the use of a penalty function that was added to the surface water treatment cost function and forces the costs of surface water to go to infinity soon after capacity is reached. When the use of surface water is beyond the point of decreasing returns or when decreas-

ing returns to dispersion dominate, we have strict concavity of the objective function, since benefits are strictly concave and all other cost functions are either linear or convex. When increasing returns dominate, strict concavity of B is not ensured, and therefore the resulting value function may not be strictly concave. The simulated value function and optimal control results for the stationary state are consistent with the above.

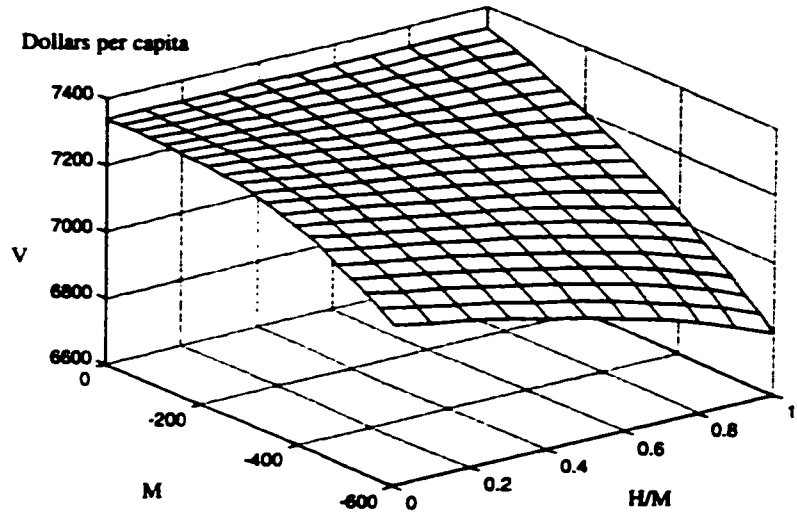
2.2 Results

This section presents the results of the numerical approximation to the end period value function, the paths of optimal controls, and the evolution of the system from 1930 to 2050. The first set of results is obtained with parameters and initial values given in table 10 page 88. The second set of results, for the same cost differential, add a linear distribution cost common to both ground and surface water, and reduces θ to obtain predicted levels of water use closer to their observed values.

A fixed point of the value function is obtained for year 2050, the beginning of the stationary state and the end of the planning horizon. Population and all spatial variables are constant after that date. A first guess of the terminal value function is obtained by calculating optimal surface water given that the use of groundwater is at the level of recharge for all state values. Optimal values of s and g and next period states are substituted into the value function, iteratively, until uniform convergence is obtained. The resulting value function (V) is smooth and concave in all dimensions, which indicates that the optimal policies obtained with this method will approximate the results of the underlying structural model. Figure 16 represents the stationary value function.

Constraints are nonbinding for all state values except at $H = -600$, where groundwater use must be at most equal to recharge. Interior solutions are obtained for all other gridpoints. Figure 17 shows the functional relation between optimal groundwater use and the values of the states, H/M and M. Optimal groundwater use increases as H increases but decreases

Figure 16: Stationary Value Function: Time Invariant State Starting in 2050



for a given H, as the difference between H and M gets smaller. A larger difference between H and M implies that more water can be withdrawn without causing any extra subsidence.

Figure 18 shows the functional relation between optimal surface water use and the values of H/M and M. Less surface water is used when water levels are higher, and more surface water is used as the difference between H and M decreases. Surface water use extends beyond the level for which diminishing returns have been imposed, which suggests that more surface water would be used if additional treatment plants could be built.

Backward induction gives us approximations of the 120 value functions for each year prior to the time invariant state. As less and less water is used, the optimization program does not run smoothly as we are in the range of increasing returns to scale in surface water treatment. The solution can jump between two local optima, one with high level of surface water use, and the other with zero surface water use. We choose two different starting values that would each reach a different local optimum in case two local optima exist. We run the

Figure 17: Optimal Groundwater Use as a Function of the States, Starting in Year 2050.

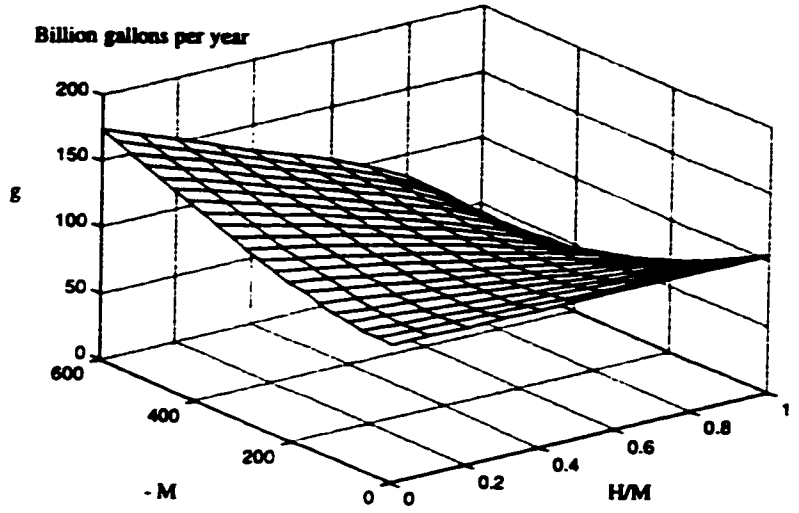
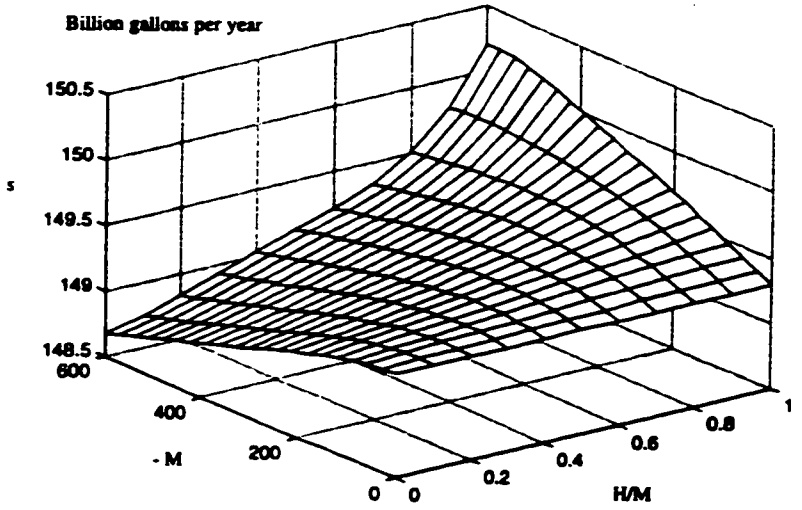


Figure 18: Optimal Surface Water Use as a Function of the States, Starting in Year 2050



optimization starting from these two starting values and select the solution which yields the largest value for the value function (the global maximum). We then use initial values to solve for the optimal paths. Figures 19, 20, and 21 show the paths followed by optimal groundwater use (g^*), optimal surface water use (s^*), and the resulting optimal path of water levels (H^*), respectively. The optimal subsidence path (M^*) follows H^* exactly since, along the optimal path, H never increases.

Figure 19: Optimal Path of Groundwater Use : Base Parameters

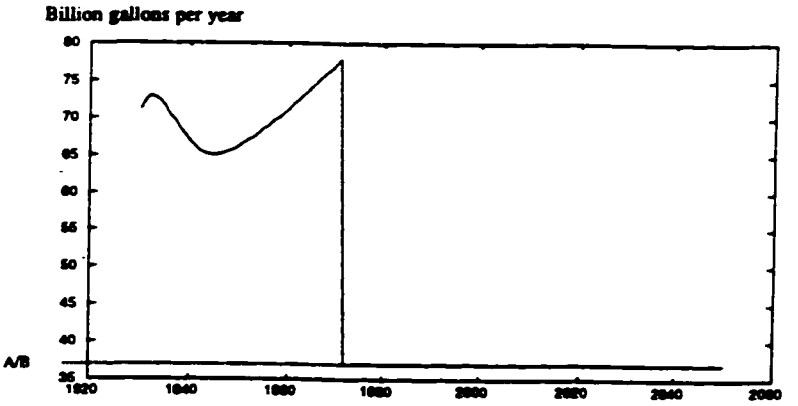
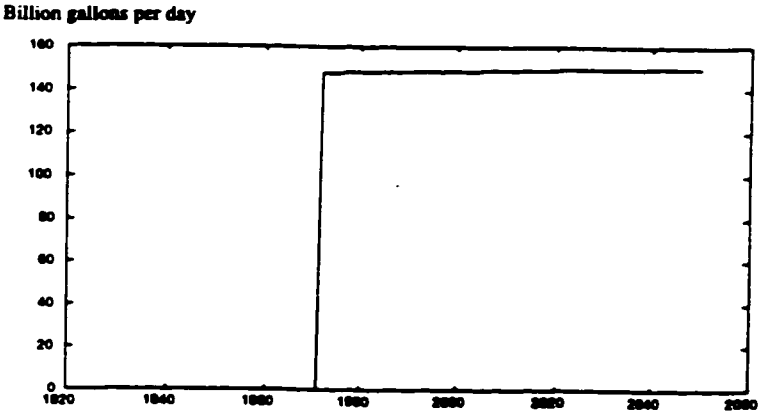
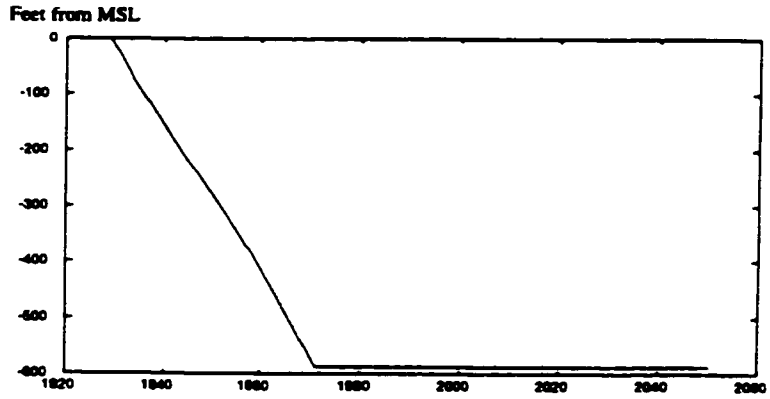


Figure 20: Optimal Path of Surface Water Use: Base Parameters



This first set of results indicate that given the level of use, groundwater will be first fully

Figure 21: Optimal Path of H and M : Base Parameters



used and depleted before surface water is used. When the bottom of the aquifer is reached, only the groundwater recharge is used and most of the rest of the water comes from the surface source. However, it is never optimal to use less than the recharge rate and to allow the aquifer to replenish, as by this time, the capacity "constraint" on treatment is reached and the marginal cost of surface water increases sharply. These results indicate that, given that the treatment plant capacity is restricted, and given a time discount of 6 percent per year, groundwater is always cheaper than surface water and it is optimal to allow full subsidence.

The level of total water use in this simulation is much higher than observed levels (at least in the first half of the planning horizon, when water use is lower and the capacity constraint on surface water is not a limiting factor). This probably occurs as we have not accounted for distribution costs, which are common to groundwater and surface water. This would increase the marginal cost of water and decrease water use. We add a cost component that increases linearly with the level of total water use. This is equivalent to reducing per capita gross benefits by $\overline{C}(s+g)/N$, where $\overline{C}$ is a constant parameter. As distribution costs represent a significant portion of total costs, the constant parameter is chosen to add 20 cents/1000 gals. to the cost of providing water in general, a reasonable estimate of

distribution costs in 1975. Adding this cost largely reduces the amount of water use in all years. However, this decrease is not large enough to reduce the simulated municipal water use in Houston to observed levels in 1955, 1975, and 1990 (22.6, 110.5, and 128.4 billion gal. respectively). Accordingly, we also modify the budget share parameter to reduce marginal benefits. With $\overline{C} = 200,000$ and $\theta = 0.002$, total simulated water use in 1975 and 1990 approximates actual water use for H and M approximately half way between their maximum and minimum levels.[92]

Since simulated water use is lower than in the first simulation, the problem of local optima needs to be dealt with in the stationary period as well as the backward induction process. The resulting value function (V) is still very smooth and concave in all dimensions, it is similar to the first run, although, as expected, the range of values of V is substantially lower. Figure 22 represents the stationary value function for $\overline{C} = 200,000$ and $\theta = 0.002$ and all other parameters as in table 10 page 88. We refer to this set of parameters as the reduced benefits case.

The difference between the two simulations shows up in the optimal values of s and g. Zero surface water use is optimal in the stationary problem when water levels are low. Figure 23 and 24 show the functional relation between the controls and the values of the states, H/M and M. It presents the same features as in the first run as long as positive use of surface water is still optimal. The variation is clear in the graph of optimal groundwater use. We see that, in the region where surface water use is positive, the variation in g is similar to the variation observed in the first run. For optimal surface water, the variation in surface water at positive values is too small to be visible in the graph. However, the graph does indicate the location of the jump from positive large values to zero optima.

Optimal paths results are presented in figures 25 (g^*), 26 (s^*), and 27 ($H^*$and M^*). Until

[92] We choose to also change θ, rather than increase $\overline{C}$, so that both parameters stay within a reasonable range.

1988, only groundwater is used. Water levels decrease steadily and subsidence continues until the water levels in the aquifer have gone down 50 feet from the "bottom" (−551 feet). Then it becomes optimal to switch completely to surface water for a period of 11 years. The use of groundwater resumes in 1999 but stays below the recharge level until 2018. During this period of no or low groundwater use, water levels recover at a decreasing rate until they reach −266 feet. After 2018, groundwater use is again above recharge, and water levels go down. In this simulation, subsidence is not allowed to resume before the end of the planning horizon. The historic low water level, which determine the amount of subsidence, remains at −551 feet.

The predicted paths of water levels and subsidence better capture the features of the analytical model described in chapter 2. During the period of water level recovery, increasing returns to scale in treatment dominate the decreasing returns to dispersion. The period of drawdown following the recovery indicate that decreasing returns to dispersion dominate

Figure 22: Stationary Value Function: Reduced Benefits Case

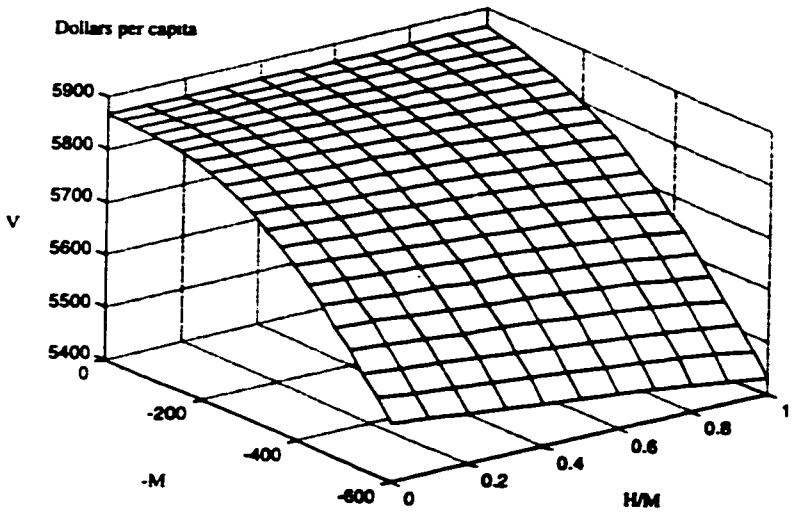


Figure 23: Optimal Groundwater as a Stationary Function of the States, Reduced Benefits Case

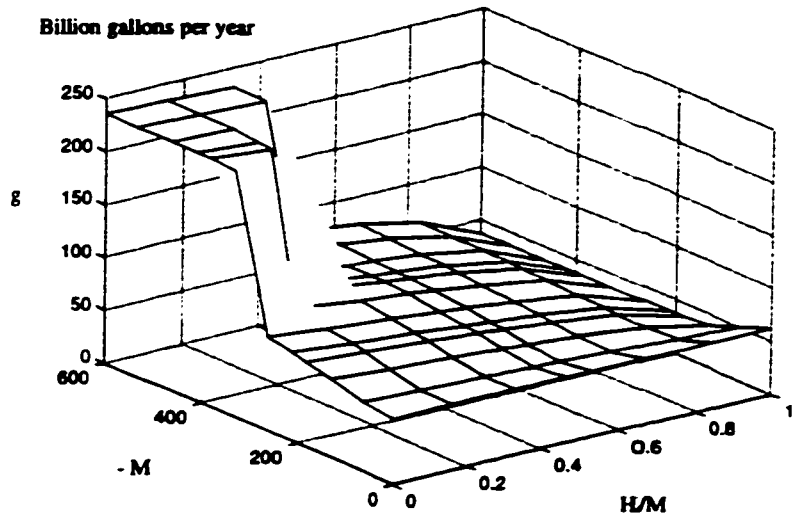


Figure 24: Optimal Surface Water Use as a Stationary Function of the States: Reduced Benefits Case

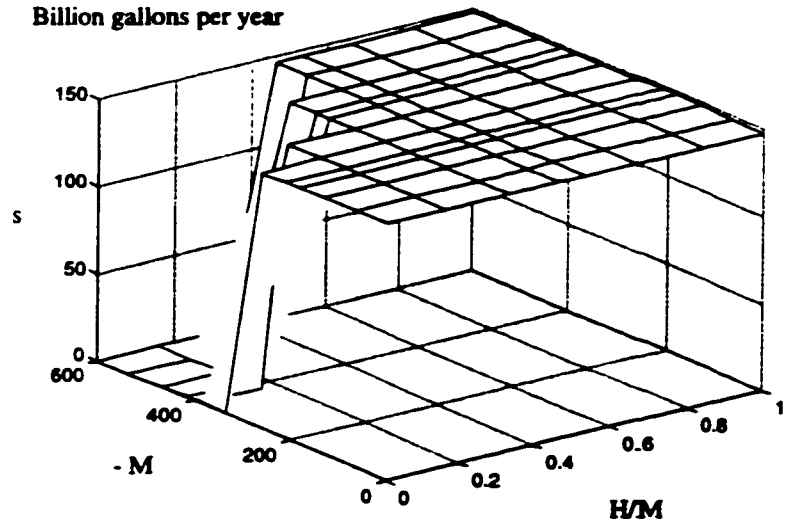


Figure 25: Optimal Groundwater Use: Reduced Benefits Case

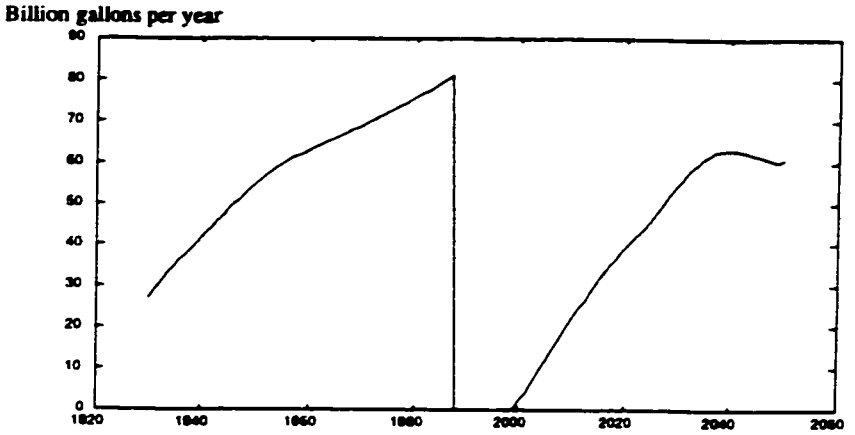


Figure 26: Optimal Surface Water Use: Reduced Benefits Case

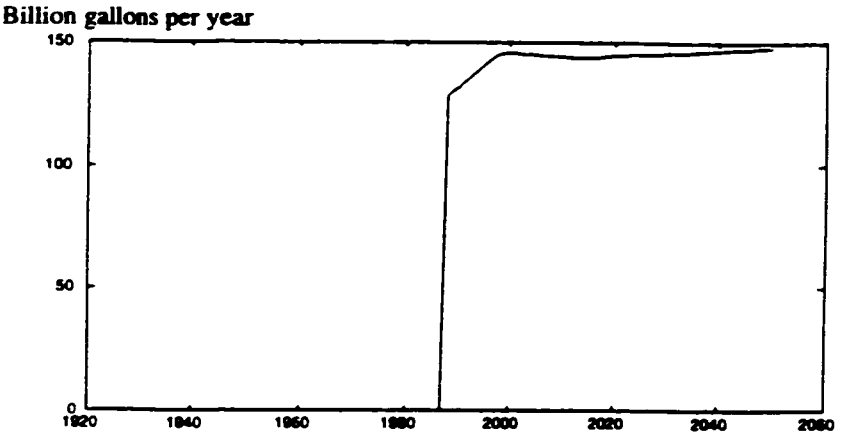
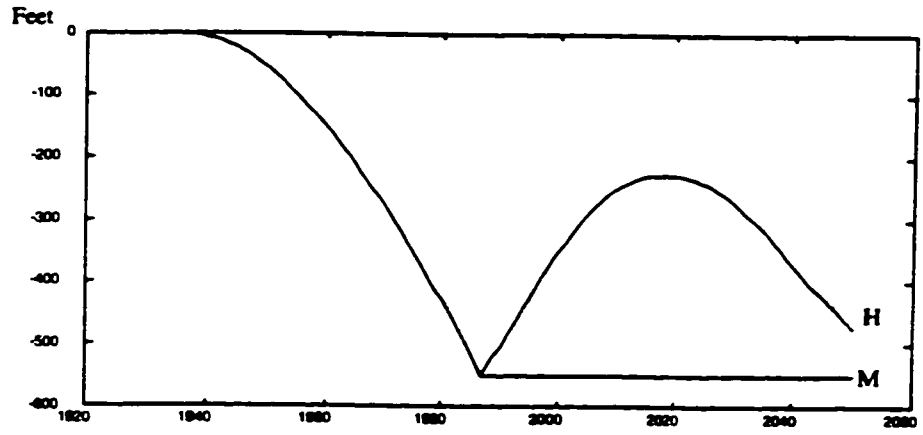


Figure 27: Optimal Paths of H and M : Reduced Benefits Case



after 2018.

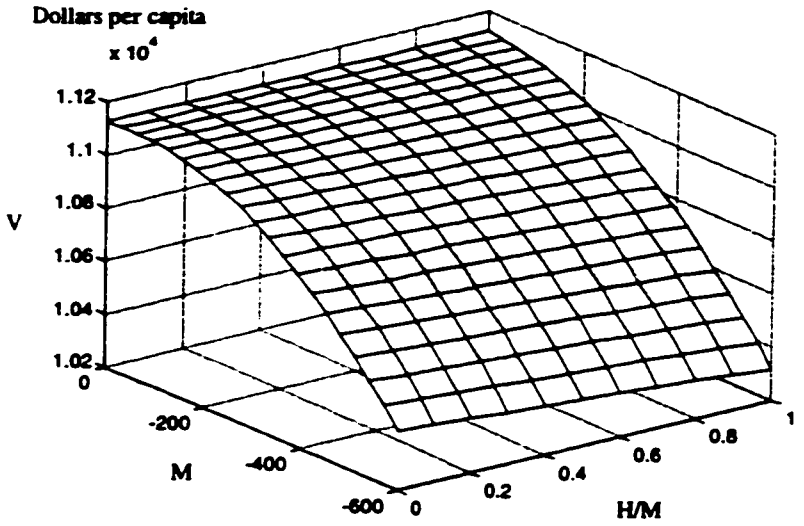
Since the reduced benefits case better approximates the history of total water use, we use this second set of simulations as a base case in carrying out various sensitivity analysis experiments. The following section examines the impacts of a lower discount rate, lower or no subsidence costs, and a different spread of the city over time.

3. Sensitivity analysis

3.1 Sensitivity to the discount factor β

The rate of discount is an important parameter determining the present value of future costs and benefits. A low discount rate (higher β) implies that future costs will have more impact on the solution. Specifically, the lower the discount rate the higher the present value of subsidence costs. We perform a new set of simulations with the rate of discount at 3 percent per year (equivalent to $\beta = 0.97$) instead of 6 percent per year. Figure 28 represents the stationary value function with all parameters equal to those of the previous simulation, but with $\beta = 0.97$. The function is very similar to the previous one but shifted up to reach dollar values twice as large as in the previous case.

Figure 28: Stationary Value Function: Low Discount Rate



The values of the optimal control variables are similar to the first simulation as they do not exhibit the discontinuities of the reduced benefits case. The use of surface water is now higher and reaches the point of decreasing returns for all values of H and M. It is never optimal to use no surface water, and so there are no jumps between local optima in the search for the stationary value function. Figures 29 and 30 represent the graphs of the optimal controls as a function of H and M. Optimal g increases at a decreasing rate as H/M decreases, for the same value of M. For low values of H/M, the difference between H and M becomes greater as M decreases (approaching -600) which explains why more groundwater is used. However, when $H = M$ groundwater remains constant at the level of recharge, except for values close to H_0 when it becomes optimal to use more groundwater (this explains the reversal in trend approaching $H = M = H_0$. The graph of s, follows an opposite pattern but always remains above 145.9 billion gallons, which is the quantity at which average costs are forced to increase sharply because the single location plant capacity

is reached.

Figure 29: Optimal Groundwater Use as a Stationary Function of the States: Low discount Rate

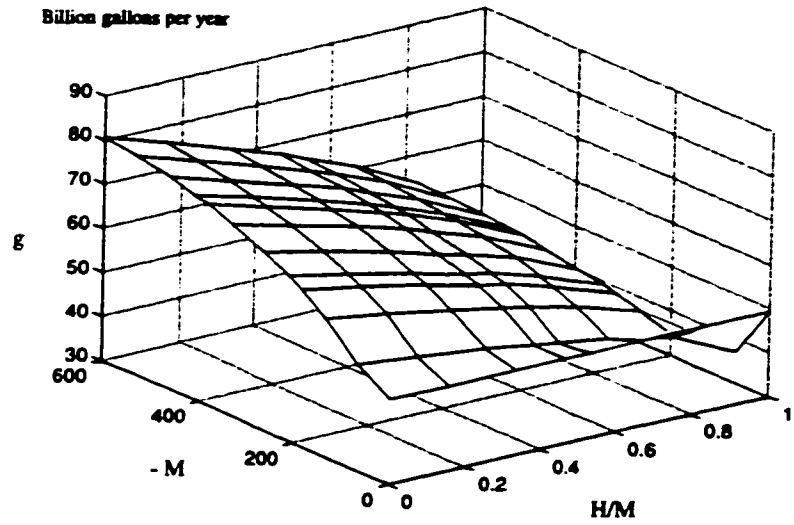
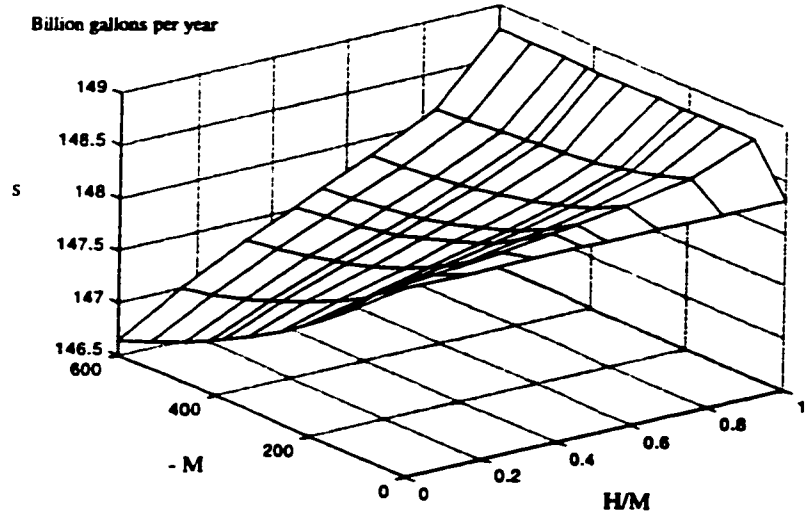


Figure 30: Optimal Surface Water Use as a Stationary Function of the States: Low Discount rate



Optimal path results are presented in figures 31 (g^*), 32 (s^*), and 33 ($H^*$and M^*). The paths are similar to the base case but reflect the fact that future groundwater subsidence costs are higher in present value terms, which justifies using less groundwater and using surface water earlier. The switch to surface water occurs earlier (1978 instead of 1988) and the period of no groundwater use extends longer (21 years instead of 11 years). Groundwater use resumes in 1998, as in the base case, and stays below recharge until 2022. Water levels follow the same down-up-down pattern, but the level of subsidence is greatly reduced. A period of sharp water level decline starts in 1938, and lasts until 1978 when the minimum water level of -327 feet is reached. Water levels then recover rapidly and reach their upper bound in 1999. The aquifer is at full capacity for a period of 23 years between 1999 and 2022. Subsidence stays at its 1978 level until the end of the planning horizon. The maximum subsidence is approximately 60 percent less than in the base case.

Figure 31: Optimal Groundwater Use: Case with Low Discount Rate

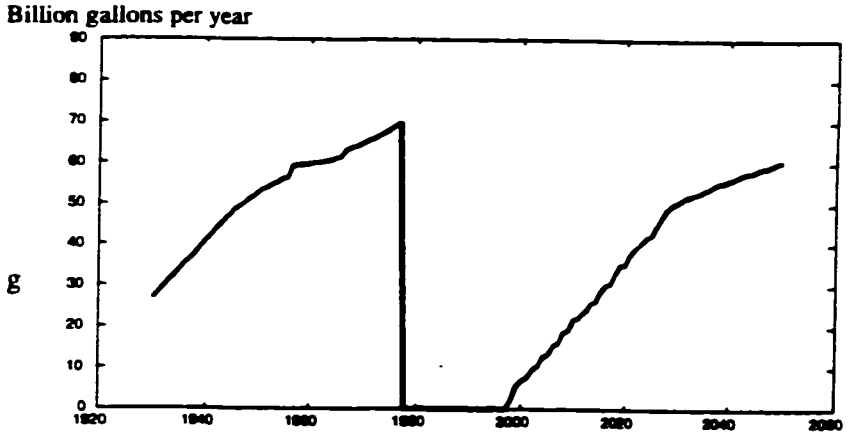


Figure 32: Optimal Surface water Use: Case with Low Discount Rate

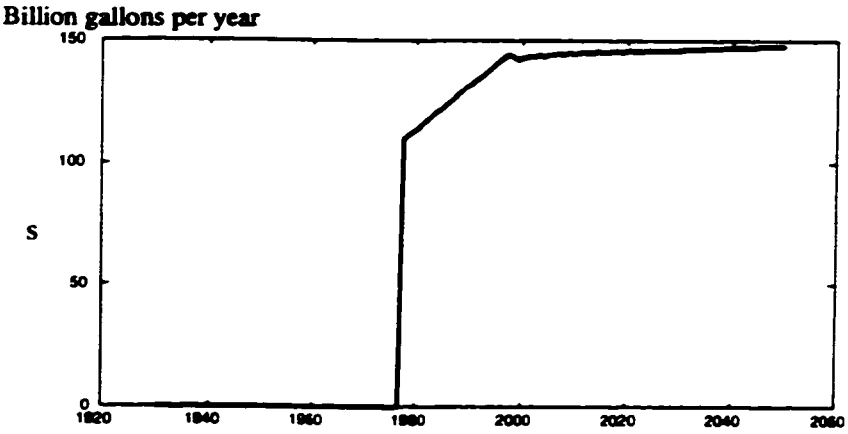
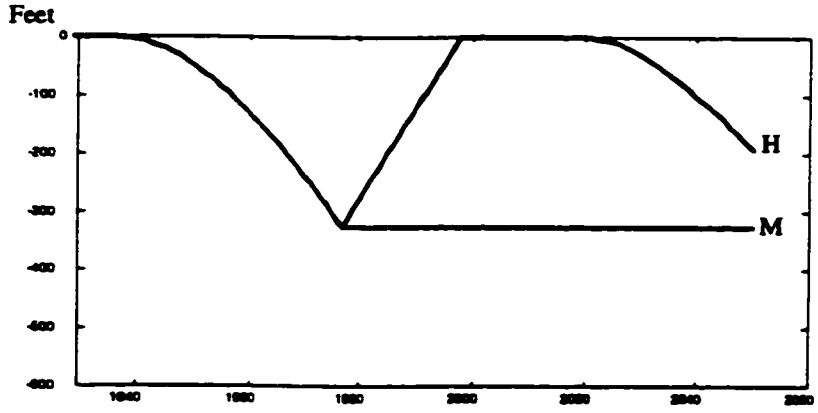


Figure 33: Optimal Paths of H and M: Case with Low Discount Rate



3.2 Sensitivity to subsidence cost parameters (values of $\overline{c}$)

Since the estimates of subsidence costs are very tentative, we need to evaluate the impact of different levels of this cost. Also, we wish to see whether high pumping costs in themselves would justify the use of surface water before the bottom of the aquifer is reached. We perform 3 simulations, one with no subsidence costs ($\overline{c} = 0$), and two with higher subsidence costs ($\overline{c} = 300$ and $\overline{c} = 400$).

3.2.1. No subsidence costs ($\overline{c} = 0$)

The absence of subsidence costs changes our results substantially. In figure 34, the value function with zero subsidence costs is compared to the base case. Figure 35 shows the new value function by itself, using a different angle to show its full range. Since there are no subsidence costs, this function shows the sole impact of the pumping externality. The value function changes almost linearly with H. This is because pumping costs increase linearly with H, but the value function is an aggregation of discounted future costs over an infinite horizon, which cause some non-linearity. As expected, the value of M has no impact on discounted net benefits. This can be seen on the graph by noticing that the vertices on the M axis and the H/M axis are identical. Along the M axis, $H/M = 1$, and therefore $H = M$ and H takes all values from -600 to H_0. Along the H/M axis, for $H/M \neq 1$, $H > M$ and H also takes on all values from -600 to H_0. Equal values for $H = M$ and $H > M$ imply that V is independent of M.

Optimal paths are reported in figures 36 (g^*), 37 (s^*), 38 (H^* and M^*). All paths follow patterns similar to the base case, but more groundwater is used in the early years and maximum subsidence is reached before any surface water is used. This indicates that the pumping externality alone does not justify any groundwater conservation. Since the full capacity of the aquifer is reached sooner, the switch to surface water also occurs earlier (1977 versus 1988 in the base case). After maximum subsidence is reached and surface water is

Figure 34: Stationary Value Function: No Subsidence Costs Versus Base Case

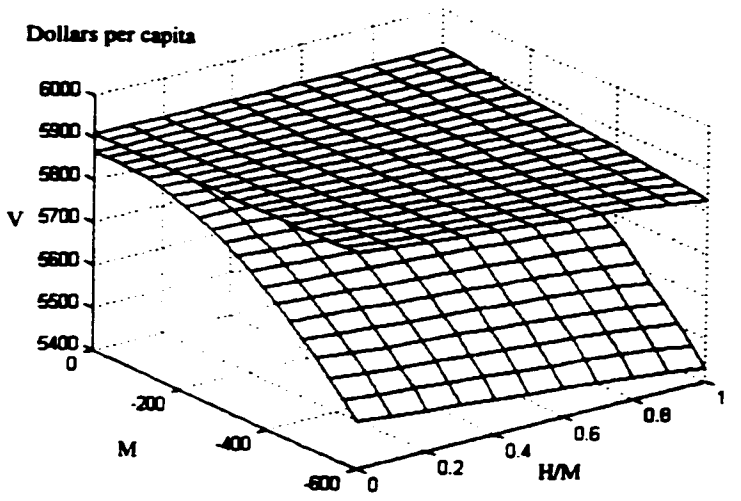
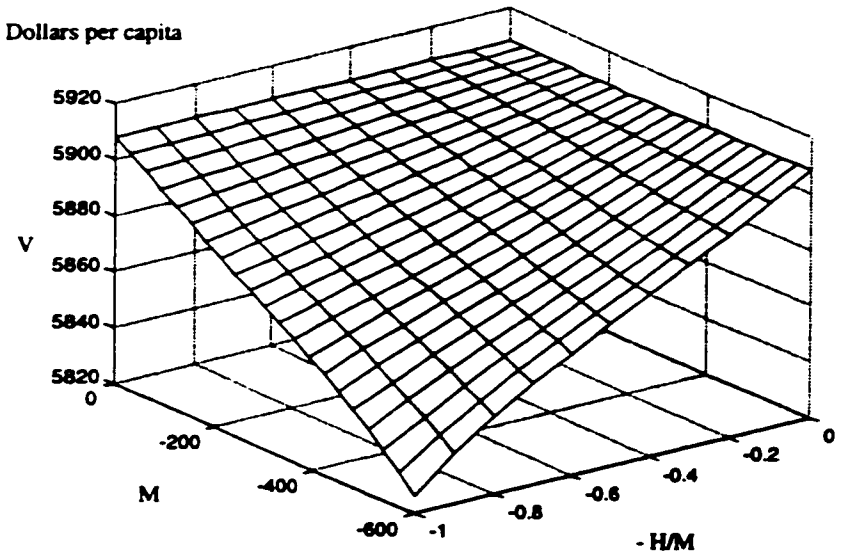


Figure 35: Stationary Value Function with No Subsidence Costs



used, subsidence costs have no impact on the solution. Therefore, as in the base case with subsidence costs, increasing returns in surface water treatment explain the recovery of water levels for a limited period of time when groundwater use is zero or below recharge.

Figure 36: Optimal Groundwater Use: Case with No Subsidence Costs

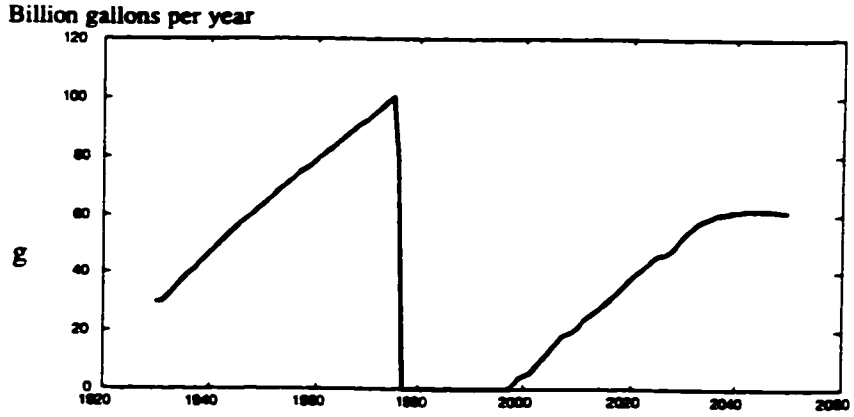


Figure 37: Optimal Surface Water Use: Case with No Subsidence Costs

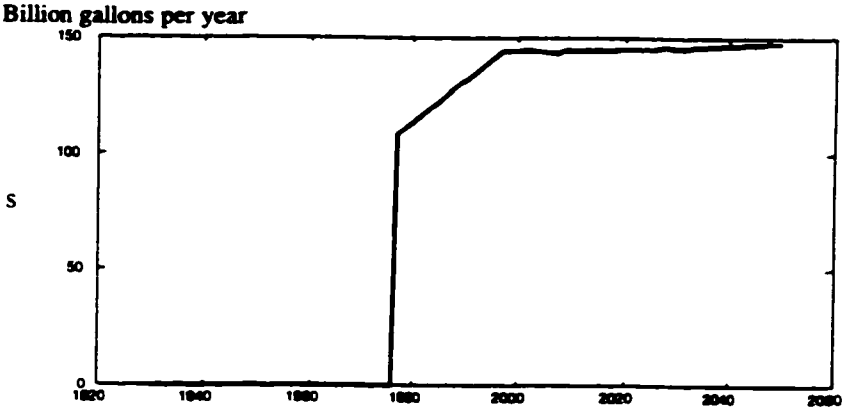
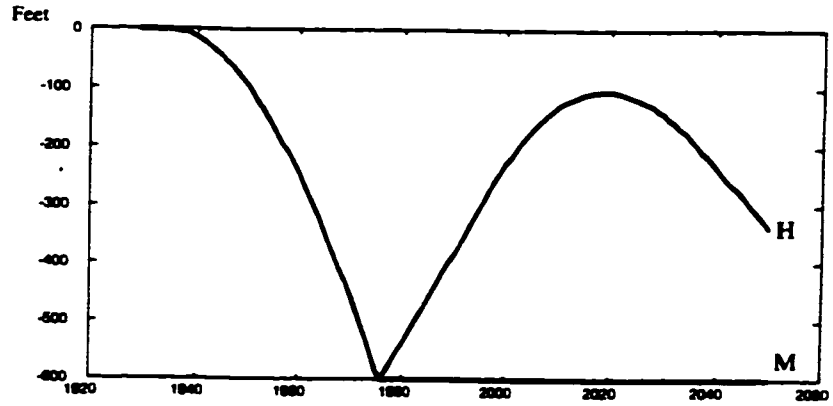


Figure 38: Optimal Paths of H and M: Case with No Subsidence



3.2.2. High subsidence costs ($\bar{c} = 300$ and $\bar{c} = 400$)

The stationary value functions have the same basic shape as the base case (see figure 22), but the larger the subsidence costs, the larger the difference between the lowest and largest value of V. For the base case of $\bar{c} = 200$, V was at its minimum of 5430 dollars per capita at $H = M = -600$ and attained its maximum of 5867 dollars per capita at $H = M = H_0$. For higher values of $\bar{c}$, the minimum and maximum of V occurs at the same values of H and M and are, respectively, 5033 and 5861 for $\bar{c} = 400$, and 5314 and 5863 for $\bar{c} = 300$.

The optimal policy variables in the time invariant state also follow a pattern similar to the base case. No surface water is used for the same values of H and M (see figure 24), and when surface water use is positive, its level is above the decreasing returns capacity level with values approximately equal to the base case. The maximum amount of groundwater use is always the same at 235.73 billion gallons per year when $H = H_0$ and $M = -600$, but groundwater use is significantly reduced when water levels and subsidence levels are low (from 62.22 billion gallons per year when $\bar{c} = 200$ to 58.32 when $\bar{c} = 300$ and 55.4 when $\bar{c} = 400$).

Optimal paths are reported in figures 39 (g^*), 40 (s^*), 41 (H^*) and 42 (M^*). The figures

show results for all three levels of subsidence costs for comparison purposes. As expected, higher subsidence costs justify an earlier recourse to surface water and lower the cumulative level of subsidence. A complete switch to surface water occurs in all cases but the time during which groundwater is not used is longer the higher the subsidence cost. This allows the aquifer to recover fully before the use of groundwater resumes. The limited planning horizon does not allow us to see whether groundwater levels ever drop below their historic low after their first recovery, which would cause renewed subsidence. However, groundwater use starts its second downward trend faster when subsidence costs are higher.

3.3 Sensitivity to the spatial parameters (l, λ, K)

Sensitivity analysis on how spatial characteristics of the city change over time allows us to examine the effect of city sprawl, a topic of concern in the urban literature in the 1990's.

The simulation was done for the case when city limits increases, the gradient decreases and density at the center decreases. The density at the center in 2050 for these parameters and the population forecasts is quite low at 2550 per square miles. A slower decrease in gradient with no change in the behavior of l so that the city limits still reach 39.3 miles in 2050 would accommodate a higher density at the center throughout the planning horizon. There are some signs in Houston that the density at the center has stopped its decline and may now increase. With $r_1 = 0.0125$, gradients are higher for every year after 1930. The gradient in 2050 would be 0.1 (the 2050 base case gradient is 0.01) and just slightly lower than the observed 1980 gradient, and the density at the center would be approximately at 10,000, which also compares to the observed 1980 level of K. Choosing this new set of density gradients will overestimate the observed gradients and central densities for Houston (for example, $r_1 = 0.0125$ with $l = 30$ implies a density gradient of 0.236 and a central density of 24,000 in 1980), but will give some indication of how concentration closer to the center would affect the solution.

Figure 39: Optimal Groundwater Use: Sensitivity to Subsidence Cost Coefficient

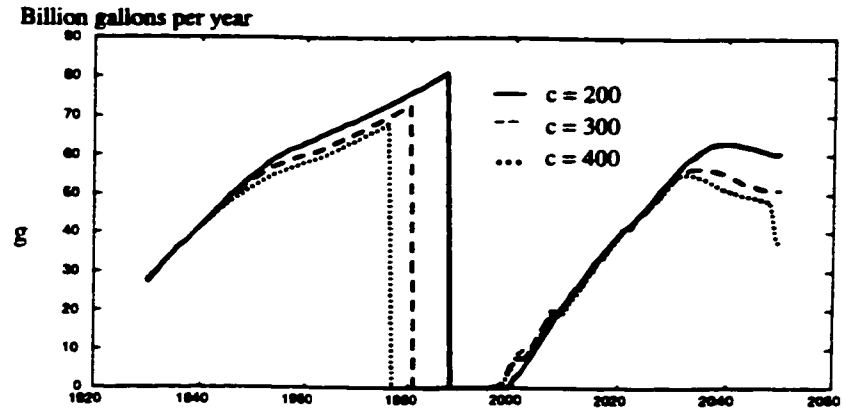


Figure 40: Optimal Surface Water Use: Sensitivity to Subsidence Cost Coefficient

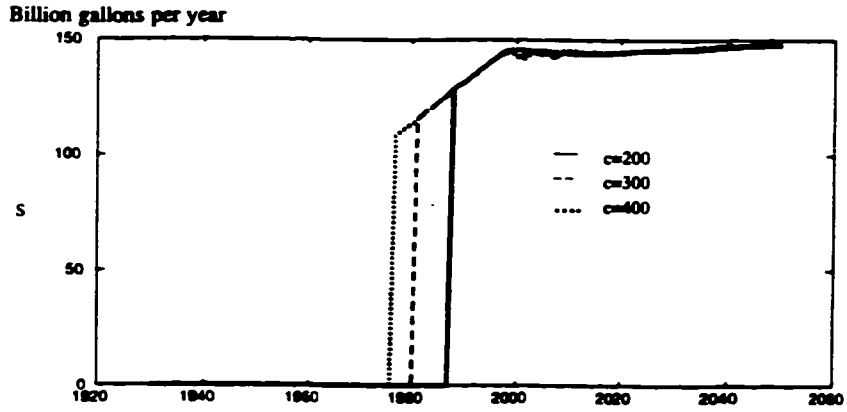


Figure 41: Optimal Path of H: Sensitivity to Subsidence Cost Coefficient

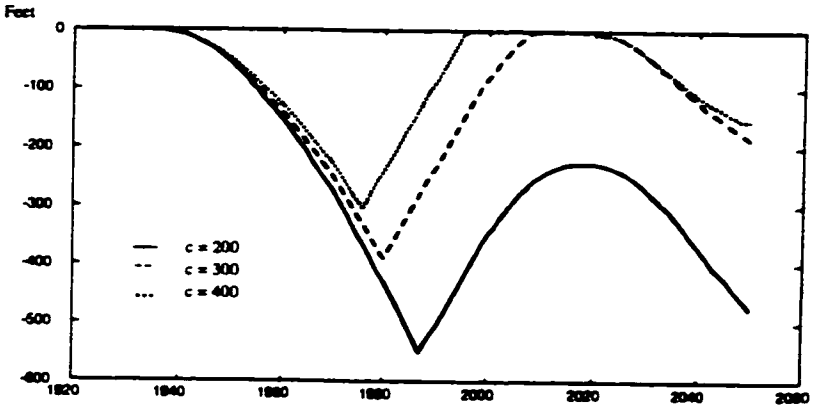
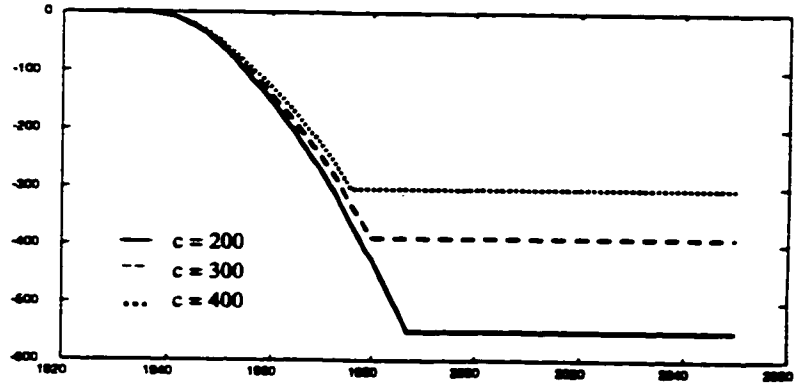


Figure 42: Optimal Path of M: Sensitivity to Subsidence Cost Coefficient



The stationary value function for the period starting at year 2050 is virtually identical to the base case with values only slightly higher (values are only 0.1 percent greater than in the base case, the maximum difference being 6.8 dollars per capita). This can be explained by the fact that, in the later years and depending on the values of H and M, either surface water is not used at all or it is used in quantities greater than the capacity of the single location treatment plant beyond which marginal costs are artificially increased sharply. Therefore, when surface water is used, surface water costs are largely dominated by treatment costs, and transmission costs become trivial. In previous years, when surface water use is below 146 billion gallons per year, transmission costs do matter and we expect some difference in the optimal paths for the policy variables and the states of the aquifer. Figures 43 (g^*), 44 (s^*), 45 ($H^*$and M^*) compare the paths of the optimal controls and states to the base case. The paths are all similar when only groundwater is used since spatial characteristics do not affect groundwater costs. It becomes optimal to use surface water at the same time and approximately when the same level of subsidence is reached. More surface water is used right away and the amount of surface water use stays higher until the capacity level of 146 billion gallons per year is reached. The difference is small (1.5 percent more surface water maximum), but since it is consistently positive, it should not be attributed to approximation errors. For the use of groundwater, the difference is also small, positive in some periods, negative in others and so, no clear cut conclusions emerge.

Optimal subsidence is not affected by the change in r_1 and optimal water levels seem to be affected only at the top of the recovery period. We expected the recovery to be quicker and H to reach a higher level, since decreasing returns to dispersion should be less important than in the base case. The graph shows a recovery slightly less than in the base case. Note that the model actually incorporate some counterbalancing effects. When more surface water is used relative to groundwater, a greater percentage of the population needs to be

reached. A lower density gradient means that the same percentage of population is reached within a smaller distance. Therefore it is possible that the optimal distance necessary to reach the population served with surface water changes too little to have a significant effect on the solution, thus explaining the low sensitivity of the results to changes in the density gradient.

Finally we run the simulation for a faster increase in city size (l increases faster if r_2 is larger). Choosing $r_2 = 1.4$ (instead of 1.2) causes the city limits to spread to nearly 50 miles from the center in 2050. Leaving the behavior of the density gradient and total population as in the base case, this implies that the density at the center is reduced for every period to about 1950 per square miles in 2050.

The stationary value function for the period starting at year 2050 is again very similar to the base case but this time, and as expected, the value are slightly lower (values are 0.1 percent lower than in the base case, the maximum difference being -5.9 dollars per capita). The same argument hold to explain why the difference is again so trivial. Figures 46 (g^*), 47 (s^*), 48 ($H^*$and M^*) compare the paths of the optimal controls and states to the base case. As water levels are increasing there is a point when it is optimal to switch back to the exclusive use of surface water. However, the use of surface water resumes right away and water levels follow their usual pattern. Again, except for the spike in groundwater use in year 2009, the paths of groundwater use and surface water use are similar to the base case. In this case, less surface water and more groundwater are used after year 2010. Subsidence is approximately the same as before but water levels do not increase as much before they start decreasing. These results are not different enough to the base case to justify concern about city sprawl.

Figure 43: Optimal Groundwater Use: Sensitivity to an Increase in Density Gradient

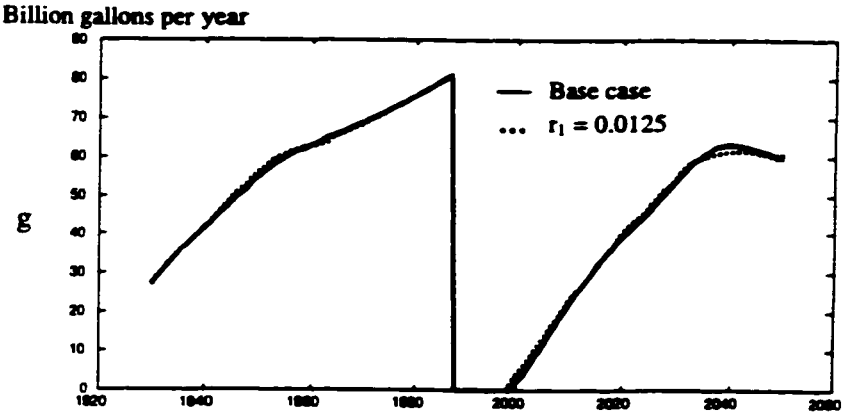


Figure 44: Optimal Surface Water Use: Sensitivity to an Increase in Density Gradient

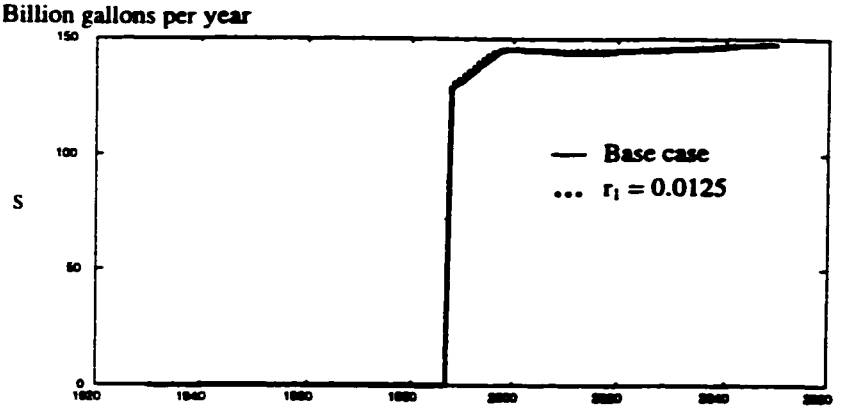


Figure 45: Optimal Paths of H and M: Sensitivity to an Increase in Density Gradient

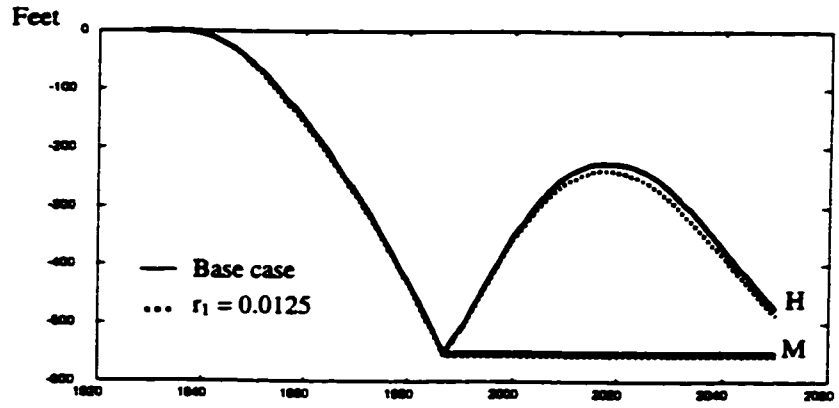


Figure 46: Optimal Groundwater Use: Sensitivity to an Increase in City limits

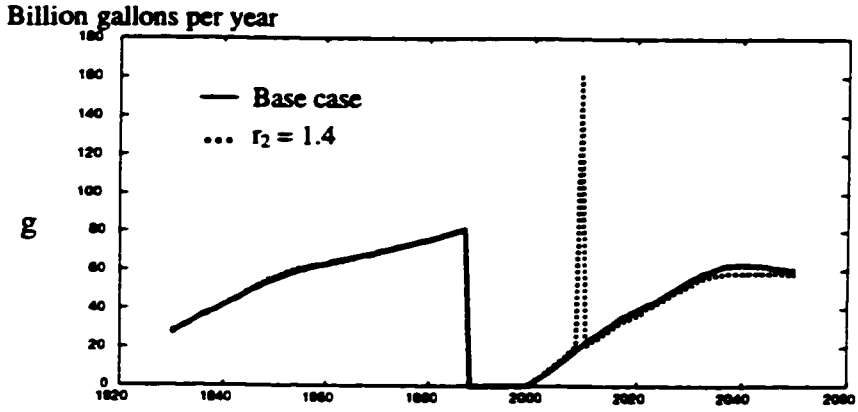


Figure 47: Optimal Surface Water Use: Sensitivity to an Increase in City Limits

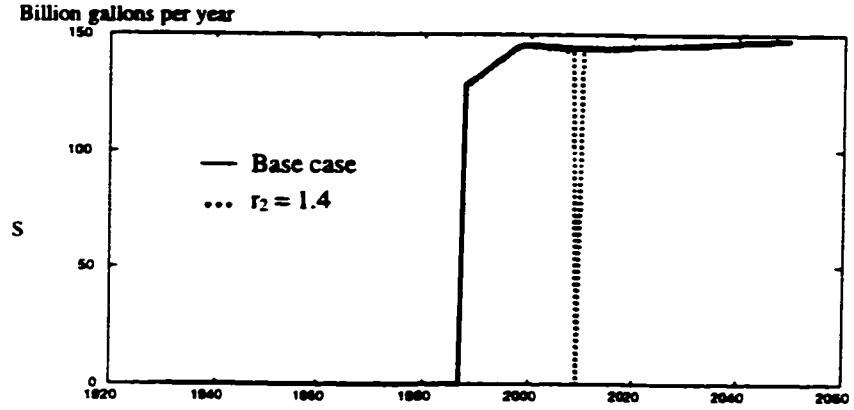
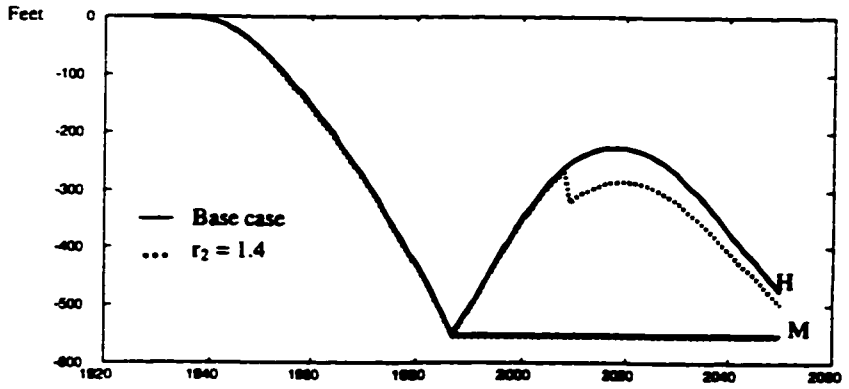


Figure 48: Optimal Paths of H and M: Sensitivity to an Increase in City Limits



CHAPTER 5

A Case Study of the Houston-Galveston Area, Texas

This chapter presents the case of Houston-Galveston, its history of groundwater use, aquifer depletion, and subsidence. The realization by city planners that subsidence was imposing a great cost on the city, following some research on the subject in the late sixties-early seventies (Gabrysch, 1969; Jones and Larson,1975; Warren et al., 1975), prompted the creation of a Subsidence District whose role was (and still is) to regulate the withdrawal of groundwater in order to halt or mitigate subsidence. The city of Houston and surroundings progressively switched to using more and more surface water. Some areas switched completely to surface water, others still use both ground and surface water. Some areas still rely entirely on groundwater. These areas are considering the costly transition to surface water to comply with the regulatory plan of the Subsidence District.

We put the facts in perspective given the results of chapters 2 and 4 and use the model to predict where the region is most likely headed in terms of groundwater and surface water use, groundwater levels, and subsidence. We then introduce the regulatory framework that is in place for the control of subsidence in the region . We critically analyze the present anti-subsidence rules in light of what we learned in the previous chapters and explore alternative policies that could be easily implemented given the legal framework in place.[93]

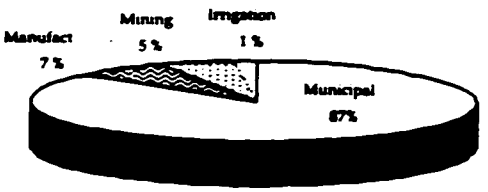
1. THE FACTS

The Houston-Galveston area draws its groundwater from two main geological forma-

[93] The Texas Water Development Board, the Houston-Galveston Subsidence District, and the City of Houston, Public Works and Engineering were very instrumental in providing me with sufficient written and oral information to write this section. Particular individuals are mentioned in my acknowledgments. I used information from interviews carried out in 1997-1998 throughout the section.

tions, the Evangeline and the Chicot. The Chicot aquifer is the smallest formation and lies above the Evangeline formation. A geologic cross section of the aquifers is provided in the Appendix (figure A1). The Evangeline aquifer is the deepest and largest and provides most of the water. It is also pumping from the Evangeline aquifer, above the natural rate of recharge, that causes subsidence in the area. Otherwise, the two aquifers are geologically similar, composed of discontinuous layers of sands and clays of about equal total thickness. The basis for distinguishing the two aquifers is mostly their difference in hydraulic conductivity. Water travels faster in the Chicot aquifer than in the Evangeline aquifer. The transmissivity of the Chicot aquifer was estimated to range from about 3000 to 50,000 square feet per day. The transmissivity of the Evangeline aquifer was estimated to range from about 3000 to 15,000 square feet per day. These aquifers span a large area so that, within the Houston-Galveston region, pumpage in one area affects water levels in other areas. The aquifer is mostly used for municipal water supply. Using data from the Texas Water Development Board we find that between 1984 and 1990, 87 percent of pumping from the Gulf Coast aquifer formation, to which the Evangeline and Chicot belong, was for municipal usage.

Figure 49: Gulf Coast Aquifer: Pumping by Category of Use, 1984-1990



Since the turn of the century, groundwater in the Houston-Galveston region has been withdrawn in ever increasing quantities to keep pace with the area's phenomenal growth. Table 11 shows the average yearly population growth rates in Harris county since 1960. From 1960 to 1996, the population in Harris county multiplied by 2.5. Population growth has decelerated each decade but is still high. This growth is expected to continue, although not exponentially (linear trend).

Table 11: Harris County Population: Average Yearly Percentage Growth

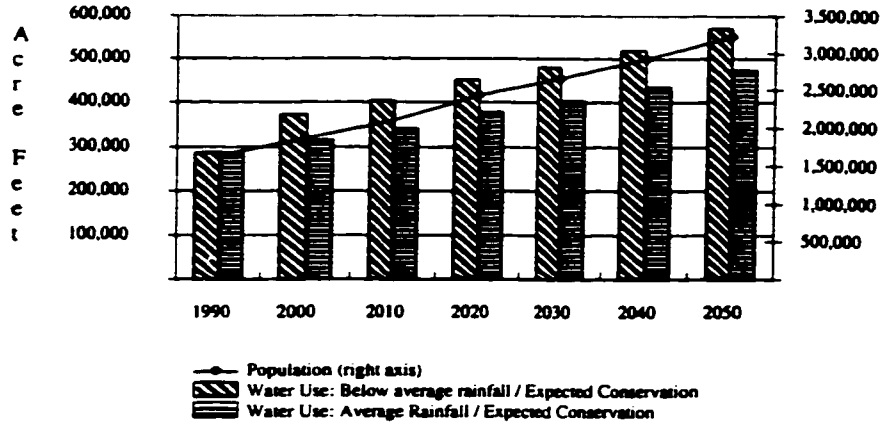
	1960-1970	1970-1980	1980-1990	1990-1996
Harris County Source: Census	3.43	3.30	1.58	1.75

Figure 50 shows the population and water use projections for Houston. Figure 51 shows the population and water demand projections for the larger Houston-Galveston area, which comprises seven counties. The Houston population is expected to be 1.63 times larger in 2030 whereas the population in the larger area including Galveston is projected to multiply by 1.84.

The early part of the century saw large increases in groundwater use in the Houston area, mostly for rice irrigation and oil refineries.[94] Groundwater use was about 11 million gallons per day (mgd) in 1906 and rapidly rose to 154 mgd in 1940. Although the growth in groundwater use was mostly for industrial needs in the 1940's, it is the rapid growth in population which became a dominant source of water demand from the 1950's and on. The city of Houston relied almost exclusively on groundwater until 1954, when water from lake Houston became available and the City of Houston East Water Purification Plant started operations with 120 mgd of capacity. This was the City's first major treatment plant. Since then both groundwater and surface water were used for municipal supply. By 1954 an aver-

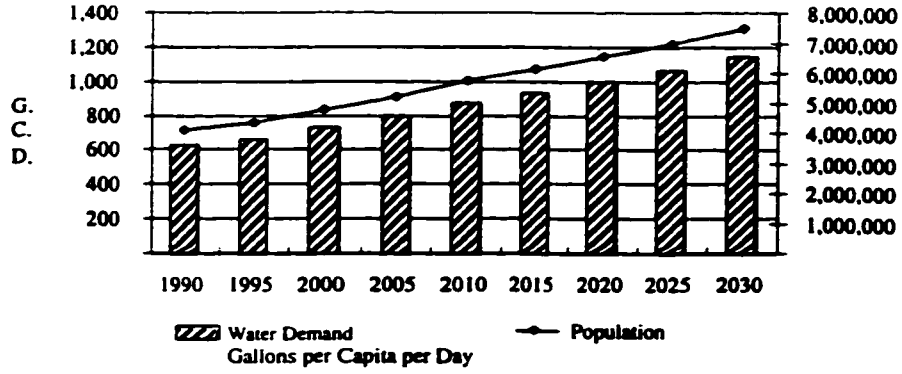
[94] The following background information on Houston water use is from an undated mimeo provided by Charles Settle of the City of Houston, Public Works and Engineering Division "Then, Now and Tomorrow: Houston's Role in Surface Water Conversion", Charles Settle, P.E. and Frederick Perrenot, P.E.

Figure 50: Houston: Water Demand and Population Projections



Source: Texas Water Development Board

Figure 51: Houston-Galveston Area: Water Demand and Population Projections

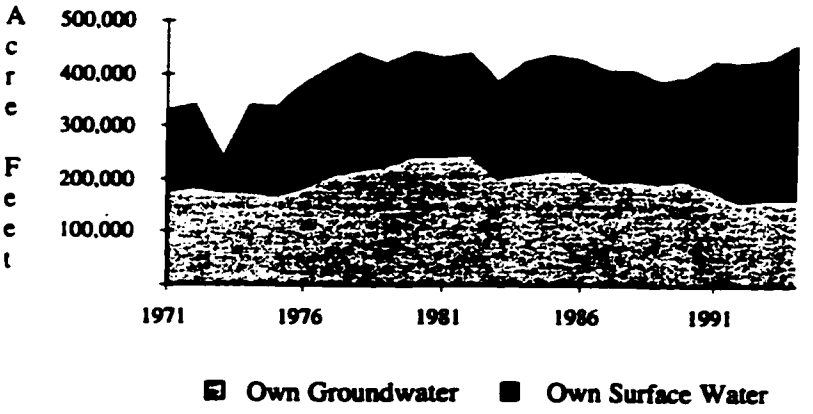


Source: Harris-Galveston Coastal Subsidence District. Report by Turner Collie &Braden Inc. March 1996

age 388 mgd was withdrawn from the Chicot and Evangeline aquifer. Large drops in water levels occurred and subsidence problems started to become significant in the Galveston Bay area. From 1954 to 1962, groundwater use was relatively stable as it increased only by 10 mgd. However, from 1962, the rapid increase resumed. By 1969, 507 mgd were withdrawn from the aquifers. This is when a new surface water source, lake Livingston, became available and a conveyance system to Houston was put in place. Since then, groundwater use has declined on average for the whole area but the spatial distribution of groundwater use has changed. While groundwater use declined in the eastern part of the region, it increase in the western part. A second treatment plant with 50 mgd of capacity started operation in 1975 next to the first East Plant. The capacity was increased in 1988 with a third plant (150 mgd). In 1990, a Southeast Plant was added, with capacity of 80 mgd. The plant is expandable to a maximum capacity of 430 mgd. Up to 1998, the city of Houston invested $125.3 millions in surface water supplies (reservoirs), $167.2 millions in treated water facilities $202.3 millions in raw water conveyance, $167.2 millions in treated water production facilities, and $183.7 millions in surface water transmission lines. With interest, this amounts to a total cost of $1.8 billion. In the meantime water bills have also greatly increased. In 1983, the bill for an average consumer using 7,000 gallons per month was $11.34, it was $20.47 in March of 1993. Water rates are still increasing. The City of Houston has now four surface water treatment plants with total capacity of 400 mgd expandable to maximum capacity of 750 mgd. Planning and design for a new treatment plant, the Northeast Water Purification Plant, started in 1995. The plant will have a capacity of 280 mgd expandable to 430 mgd. A map of the surface water treatment plants with capacity description as well as all the surface water supply sources and water rights is provided in the Appendix, figure A2. Despite this dramatic increase in surface water supply, groundwater still constitutes more than 50 percent of total water use for the city. Figure ?? shows the amounts of groundwater

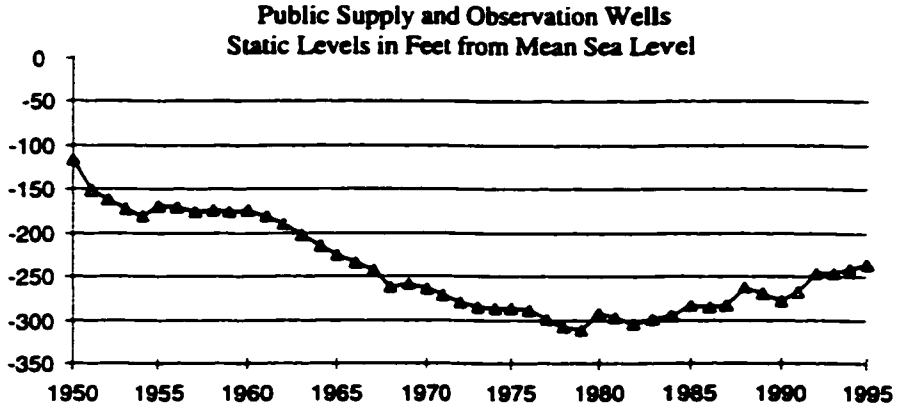
and surface water use in the city since 1971. Around Houston, nearly all the subsidence is due to groundwater pumping and subsidence is directly proportional to the amount of water derived from the clay beds of the aquifers, which corresponds to16 to 31 percent of total pumpage (Carr et al., 1985). Figure 53 shows the average change in water levels in the Evangeline aquifer since 1940. Figure 54 shows the relationship between water level in the aquifer and water use in Houston.

Figure 52: Houston, Water Use by Source



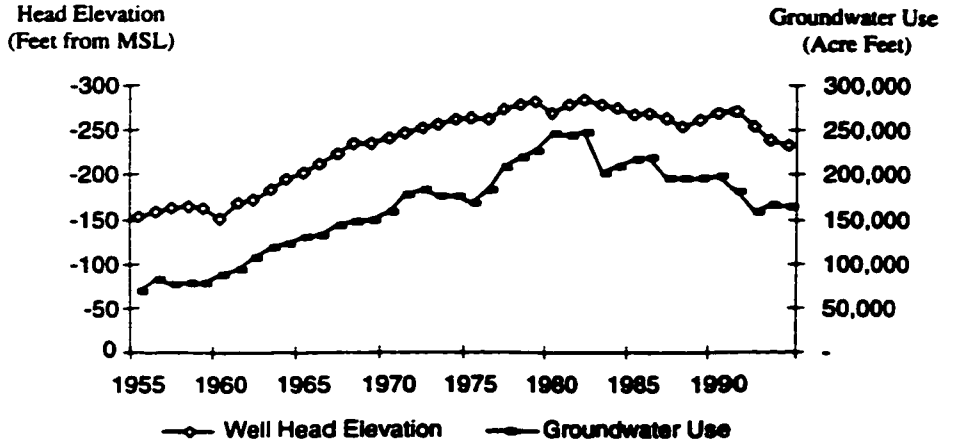
Subsidence accelerated with increased pumping, rendering the flat terrain more susceptible to flooding, inundation, and flood surges. Tragic examples of the devastating effects of subsidence are apparent in the Brownwood subdivision in Baytown. Many homes, once high and dry, are now sitting in the Bay. The area of greatest subsidence occurred in the vicinity of the Houston Ship Channel between 1906 and 1978 (up to approximately nine feet of subsidence since 1906). The Pasadena area, in the Eastern part of the region suffered the greatest amount of subsidence at about 10 feet. Costs of subsidence tend to be more acute in coastal areas because of flooding. In non-coastal areas, however, faults specific to the region are activated, worsening the impact on structures. Total annual costs of

Figure 53: Average Elevation of Well Heads, Evangeline Aquifer



Source: Texas wells database, Texas Water Development Board

Figure 54: Relationship Between Groundwater Levels and Groundwater Use in Houston



subsidence during the 1969-1973 period were estimated at $31.7 million per year. Given such an estimate of the external cost, surface water is a much cheaper alternative in the long run (Jones and Larson, 1975).

Since 1976 (which corresponds to the time when the second East water treatment plant became operational and the Houston-Galveston Subsidence District was created), subsidence has been halted in the most critical coastal areas as municipalities and industries converted to surface water.[95] The subsidence problem has now moved inland where rates of subsidence are still increasing. Before turning to the regulatory issues dealing with the Subsidence District' s plans against subsidence, we first consider the facts in the light of the model developed in the previous chapters.

2. Model Predictions and the Case of Houston-Galveston.

The results of chapters 2 and 4 can provide some insights about the facts laid out in the previous section. Houston-Galveston can be put in perspective in terms of where the region most likely stands in the different stages of our model. In particular, given that water levels are going up on average, the Houston-Galveston area must be in the phase of development where some areas are taking advantage of increasing returns to scale in surface water treatment and this effect dominates. Many Houston suburbs have completely switched to surface water, taking advantage of the existing water treatment plants (Southeast and Northeast), or building their own. The areas closest to the treatment plant use mostly surface water, while the areas furthest from the plants (Northwest and Southwest, still rely heavily on groundwater. In general, the average rate of subsidence has declined significantly.

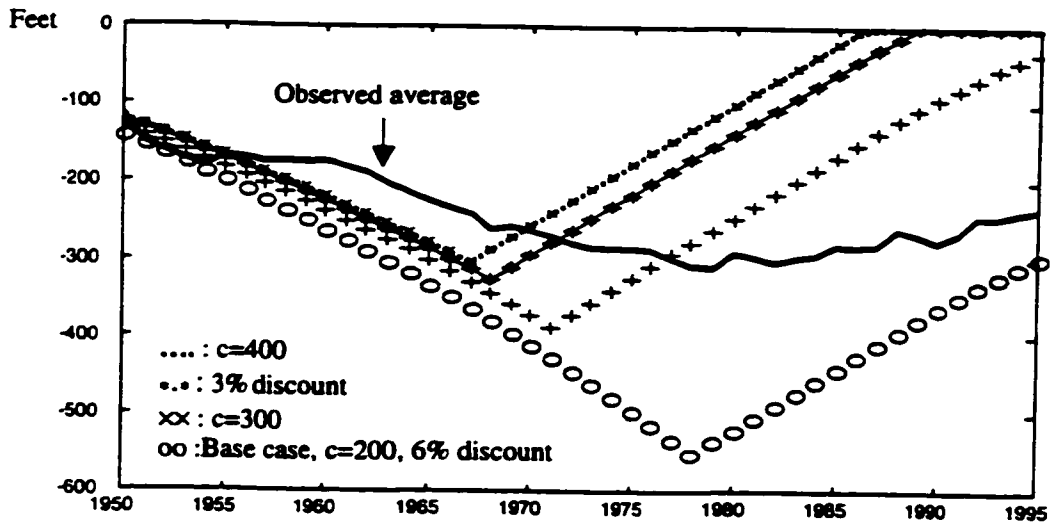
Figure 55 plots the simulated path of water levels for the base case of chapter 4 and for either a low discount rate or high subsidence costs, and compares them to observed

[95] Figures A2, A3, and A4 in the Appendix show the amount of subsidence in the region and its location from 1906 to 1995, from 1973 to 1995, and from 1987 to 1995.

average water levels from figure 53. The average level of subsidence observed is compatible with a low discount rate or the highest value for the subsidence cost parameter ($\overline{c} = 400$). However, in both simulations, maximum subsidence is reached earlier than observed. The base case reaches maximum subsidence in the late seventies as observed, but the simulated level of subsidence is higher. The simulation results indicate that Houston did enough to mitigate subsidence and perhaps too much. A longer initial period of groundwater use while a larger treatment plant was being built, would have increase subsidence but would have decreased costs by more than the present value of the extra subsidence costs. The use of surface water in Houston started earlier than in any of our simulations, and so the decrease in water levels is slower in the initial period of water level declines than for all the simulated paths. This is explained by regional disparities in subsidence levels and the fact that we rely on averages. The Baytown and Pasadena areas had experienced a much higher level of subsidence than other areas to the north and north-west. Also, Houston did not build the full single location treatment capacity right away and added capacity only progressively. This created an added constraint on the use of surface water (full scale economies could not be reached), explaining the more gradual increase in surface water, the continued use or groundwater, and the smoother path of actual water levels relative to simulations.

The population projections for Houston and surrounding areas indicate that Houston's population is expected to grow more slowly than the population in the surrounding areas. Taking Houston to be the central district, this indicates a further decrease in the regional population gradient. The predictions of our model indicate that decreasing returns to dispersion will eventually dominate and there should be a phase when groundwater levels will go down again before new treatment plants are put in operation. Since the resumed decline of groundwater levels is likely to worry city planners, it is important to understand that decreasing water levels naturally follow an increase.

Figure 55: Comparation of Actual and Simulated Paths of H



The pattern of building treatment plants is consistent with the model. First, capacity was increased by building new plants (extra capacity) in the same location. This occurred until 1990, then Houston started to build new plants at different location (southeastern in 1990 and a northern plant that should be starting operations soon). This should corresponds to a period when decreasing returns to dispersion are anticipated to dominate and it becomes necessary to diversify treatment plant location. In the simulation, the water level increase lasts a long time and further declines do not occur until after year 2015.

3. Regulation Against Subsidence

3.1 The Harris-Galveston Coastal Subsidence District

In May 1975, the HGCSD was created by the 64th Texas legislature by passage of special statute House Bill (HB) 552 in response to groundwater depletion and large subsidence problems. The HGCSD is a special case of an Underground Water Conservation District (UWCD). Chapter 52 of the Water Code normally provides the board of directors

of a UWCD with powers including:

-research and planning

-making and enforcing rules

-requiring permits, records, reports and drilling logs for wells

-regulating spacing and production to minimize drawdown and subsidence

-filing out for injunctive relief in district court for violation of its rules and regulations

- levying ad-valorem taxes and issuing bonds

-exercising eminent domain to construct dams or to drain lakes

A UWCD cannot, however, sell or distribute either groundwater or surface water.

The Act creating the HGCSD contains some provisions of Chapter 52 but its administrative provisions and enforcement powers are different. Unlike Chapter 52 for UWCDs, no confirmation election was required before the District could begin operating. Moreover, its 15 member board is appointed rather than elected. The HGCSD was to be financed by permit fees charged to well owners coupled with ad-valorem tax revenues. The use of ad-valorem taxes was ruled out before the District could be approved. Therefore, the District has no authority to sell water or raise taxes. Funds generated by the permit fees are earmarked to finance the cost incurred by the District in issuing groundwater permits and performing its regulatory functions.

The purpose of the HGCSD is to regulate groundwater withdrawals within the boundaries of the District in an effort to halt subsidence "which precipitates flooding, inundation, or overflow of any area within the district, including without limitation rising waters resulting from storms or hurricanes."

One section of the Bill enjoins the Board to exercise its authority in a reasonable manner: "...taking into account all factors including availability of surface water, economic impact upon persons and the community, degree and effect of subsidence upon the surface of land,

and differing topographical and geophysical characteristics of land areas within the district."

The HGCSD considers among the relevant factors for determining whether to issue a permit or limit groundwater withdrawal the "quality, quantity and availability of surface water at prices competitive with those charged by suppliers of surface water within the District". It also includes as a relevant factor "an equitable distribution of available groundwater" (HGCSD, 1985, p. 5).

3.2 The District Plans

Aside from setting the amount of fees to be collected, the HGCSD has a role in planning and restricting future water use. This role is quite important given the rapid increase in population in the Houston area.

The District adopted in 1976 an interim plan for the areas most affected by ground water subsidence, namely the southeastern part of Harris County and all of Galveston County, by designating the area as an "Area of Concentrated Emphasis" or ACE. Subsidence was successfully halted in this area by the early 80's. Two factors contributed to the success: (1) the availability of surface water, and (2) the conservation of water by industry. While the District was focusing on the ACE, extensive growth was experienced in western Harris County. As a result, groundwater withdrawals have increased dramatically causing greater subsidence inland.

In 1985, the District published a plan to extend the focus to inland areas. The District argued that, "although inland areas are not at risk from storm surge and coastal flooding, groundwater withdrawals in those areas affect subsidence not only in the inland areas but also in the coastal areas". The District emphasizes the fact that pumping in one area affects subsidence in other areas, specifically that inland overdraft will eventually affect coastal subsidence even if no water is pumped from the coastal areas. The District was now better informed and had the ability to use simulations from a new computer model which was able

to predict rates and amounts of subsidence given projected population growth (Carr et al., 1985). It could also project the amount of surface water use necessary to meet the objectives of the Plan.

The 1985 "Regulatory Action Plan" divides the District into eight areas and establishes requirements for each area (see Appendix, figure A6). It prescribes maximum ratios of groundwater withdrawals to surface water usage and specific years in which the target must be met. In general, between and after deadline years, increases in groundwater withdrawals may be permitted provided it does not substitute for surface water use. The plan covers a horizon until 2020.

Area 1: the original ACE. No increase in groundwater withdrawal is permitted through 1989. Beginning in 1990, groundwater use must be reduced so that no more than 10 percent of the total use is from groundwater.

Area 2: Groundwater must constitute a maximum of 20 percent of water use in 1990, 1999, 2007, 2015.

Area 3: Groundwater must constitute a maximum of 20 percent of water use in 1995, 2012.

Area 4: Groundwater must constitute a maximum of 20 percent of water use in 2000.

Area 5: Groundwater must constitute a maximum of 20 percent of water use in 2000.

Area 6: Groundwater must constitute a maximum of 20 percent of water use in 2005.

Area 7: Groundwater must constitute a maximum of 20 percent of water use in 2010.

Area 8: Increases in groundwater withdrawals may be permitted but groundwater shall not be supplied to areas outside the boundaries of area 8.

Aside from the coastal areas, the different dates are calculated so that the area has time to realistically introduce additional surface water given the time and cost of constructing surface water treatment facilities and transmission pipes to the existing distribution systems.

In 1992, new technical studies and the recognition of a change in growth patterns in the Houston area allegedly prompted the HGCSD to change the number and location of regulatory areas (see Appendix, figure A7). The focus for earlier conversions was shifted from Southwest to northwest Harris County. Regulatory area 3 was substantially reduced in size but retained the same 1995 conversion date.

4. Analysis of the HGSD Regulation-Spatial Issues

The first issue to be addressed is whether it was wise to create a District unable to raise taxes or sell. A simple tax per gallon of water extracted from the aquifer could have been enough to re-establish the correct price signal for groundwater and make surface water a least cost alternative. In California, the Santa Clara Valley Water District managed to halt subsidence by imposing surcharges to pumpage in zones where groundwater drafts were occurring. Orange County, California, has also been operating successfully under similar rules (Teutsch, 1978). I choose to leave this issue aside since the provisions of the Act creating the HGCSD do not give the powers to the District to levy such taxes. Instead, the HGSD must rely purely on regulation and fines. Since the entities targeted are all public entities, such as municipal utility districts, the HGCSD plans are actually taken seriously, and deadlines are anticipated accordingly through planning. In effect, the strategy of the HGSD is to decide how much groundwater is safe to withdraw in order to mitigate and eventually halt subsidence, and to set quotas to the different areas so that the goals can be met. I focus on the HGCSD Plans outlined in the previous section. I show that municipal utility districts seeking to minimize long term costs of compliance have actually been instrumental in modifying the plan to make it more cost efficient. However, the plan is still not recognizing that a less equal share of the safe groundwater yield may actually be more efficient to the entire area.

4.1 The 1992 Changes are Consistent with a Better Cost-Minimizing Path

Why was the 1985 plan modified? Was it really due to new population growth prospects and new technical information? The shift of focus from southwest and west of Houston to the north and northwest does not seem compatible with changes in patterns of population growth. The latest subsidence maps (Appendix, figure A3) seems to show the highest subsidence in the new area 6, which includes the old area 4 and area 7 (Appendix, figures A4 and A5). The deadline for conversion did become shorter for some for the part of area 6 that was area 7 from 2010 to 2005 but is still later than the new area 4. The explanation seems to lie elsewhere.

In 1988, the North Harris County Water Supply Corporation (NHCWSC, a group of Municipal Utility Districts or "Super MUD") adopted a plan for switching its water supply to a surface water source (NHCWSC, 1987). Until then the area was using groundwater exclusively at an average variable cost of $.25 per 1000 gallons. The lowest cost alternative for surface water provision was going to increase that cost to $.75. The capital costs were to be paid by issuing bonds to be repaid using an increase in property taxes. The chosen alternative involved using water from lake Houston. The NHCWSC was to share the cost of the new Northeast water plant, which the City of Houston had already decided to build, up to a capacity of 10 mgd. The cost of transmission lines from the plant was to be shared by different localities along the line depending on distance and capacity.

Did the HGCSD know of that plan and decide to change the deadline for this area to an earlier date because it would have met it anyway? It seems plausible. But why would areas 4 and 5 be changed to include the whole Northeastern part of the District?

The construction costs of the plants are usually assumed constant at about 1.20 million dollars per gallon of daily capacity (NHCWSC, 1987), but the average cost of treating water goes down considerably with size. In terms of the conveyance system from the plant,

there are some economies of scale in constructing larger new pipes. One large pipe is less expensive in construction and operation than two small pipes carrying the same total capacity. Thus, economies are to be expected when adjacent areas plan to switch to surface water in a coordinated fashion. However, the cost of transmission is an increasing function of distance. As Dajani and Gemmel (1973) pointed in their Guidelines for Public Utilities Planning and consistently with the of the assumptions formulated in our previous analysis, "Economies of scale may be offset by diseconomies of dispersion, agglomeration, or spatial arrangement and pattern".

The 1985 plan did not take these considerations into account. Areas 4 and 5 have the same deadline of 2000. They are located at the opposite sides of the county. Area 5 is close to surface water sources and to area 2, which has already switched to surface water. Area 6 is also in a better position than area 4. The basis for deciding on the deadlines in 1985 did not seem to include the role of economies of scale in treatment and the possibility of building larger transmission pipes and share their cost when service areas are not too far apart. The 1992 changes in the layout of zones are consistent with a better utilization of potential economies of scale. The new zones are consistent with the location of surface water supply sources and planned water treatment plants. The new area 4 includes the whole area of the NHCWSC and goes all the way to lake Houston, where the new water treatment plant is under construction (see Appendix figures A6 and A7). The NHCWSC showed good foresight and was quick to recognize that it could take advantage of its geographical location. The flexibility of the HGCSD in changing the zones in response to the new information is remarkable.

4.2 Case against Equal Use of Groundwater: the Tradable Permit Alternative.

The discussion above points to the fact that some municipalities in the western and southwestern parts of Harris County would have an extremely high cost in transmission if they

were to use the same surface water supply sources as the Northern and Eastern areas. Why did the district insist on having an equitable distribution of the available groundwater? (Available amount of groundwater refers here to the safe yield or equivalently the quantity that can be withdrawn without causing depletion). Each area can use up to 20 percent of all their water from groundwater. If the hydrology of the aquifer is such that two areas are well interconnected so that reduction of pumpage in one area permits increase in pumpage in another area while keeping the overall water level the same, then there would be some benefits to unequal use of the safe yield. If an area is close to a raw source of water and a treatment plant, it could rely completely on surface water, allowing one of the remote areas to rely on more groundwater. This of course can only be done equitably if the remote area is compensating them for the extra cost of using surface water.

The problem of allocating a limited amount of groundwater to an area in which costs of alternative sources differ is similar to the problem of allocating limited amount of pollution rights, for which the benefits of tradable permits was clearly established in the literature (Stavins, 1998; Crocker, 1966; Dales, 1968; Montgomery, 1972). When groundwater is not mined, the external cost of subsidence is zero. The private cost of groundwater is approximately equal to its social cost. The safe yield of the aquifer should therefore not be wasted and it seems reasonable to allocate its use equitably among the different areas in the district. However, if areas are sufficiently hydrologically interconnected, then the different areas should be able to trade their permits among themselves so that areas with lower surface water costs would use less or no groundwater and areas with high surface water costs would be able to still rely on groundwater. The area with low surface water cost would prefer to sell its permit to another area if it can receive an amount greater than its cost of providing surface water instead of groundwater. The area with relatively high surface water cost would prefer to buy the permit if it is cheaper than the surface water supply alternative. The

total cost of reducing the use of groundwater to its safe yield would therefore be reduced. The Subsidence District justifies the regulation of the entire region by claiming that inland withdrawals affect conditions in the coastal areas. Given the approximate transmissivity of the Evangeline aquifer estimated to range from about 3000 square feet per day to 15,000 square feet per day, this would indicate that water travels less than a mile in a month. If these estimates are accurate, this would undermine the validity of the permit alternative in the short run because costs of pumping would not decrease right away for the area relying on groundwater. Some scheme of deferred payment would be necessary.. Nonetheless, relying on the Subsidence district's assumptions, a system of permits would have the potential to achieve the desired goals at a lesser total cost, and would be more equitable cost-wise to the different areas. Given the provisions of House Bill 552, it seems that such a plan would be consistent with the mission and regulatory powers of the HGCSD.

Many interesting questions remain about the feasibility and implementation of such permits. Once the tradable permit alternative has been accepted as a reasonable alternative, the implementation of such a plan needs to be examined in more detail. For example, how should permits be initially allocated to different areas? Should the allocation of permits be dependent on the current amount of groundwater used? What mechanisms need to be in place to enforce these permits? If there are parties outside of the regulated area drawing on the same groundwater supply, how should they be handled? These questions are beyond the scope of this dissertation but should motivate further research.

General Summary and Conclusions

Decisions over the provision of water to residential customers in large municipalities often involve choosing between two sources of water: surface water and groundwater. The two sources have different direct costs and different environmental impacts. Ignoring pumping externalities and environmental impacts, groundwater is commonly a cheaper alternative for the provision of residential water. Indeed, several factors work to increase the cost of surface water relative to groundwater. First, surface water production is centralized whereas groundwater lies below the surface and can be tapped with wells in many different locations. Surface water is more expensive because of higher transmission costs. Second, extensive treatment is necessary to supply surface water that meets quality standards, whereas uncontaminated groundwater requires minimum treatment at the well head, further increasing the cost of surface water relative to groundwater. However, the use of groundwater causes inter-temporal externalities when water levels in the aquifer decrease as a result of pumping quantities greater than the recharge (mining). In most cases, the pumping externality alone would not justify using surface water and the aquifer would be drawn down before surface water would be used. The presence of environmental costs associated with the mining of aquifers is potentially a much more significant factor in the decision of cities to consider surface water as an alternative source. This thesis examined the role of land subsidence on the solution to the optimal inter-temporal use of groundwater in cities where surface water is a feasible alternative. The model was set up as a social planning problem in which a city planner chooses optimal amounts of ground and surface water for all future periods so as to maximize the present value of net benefits to the city. Population was assumed to increase linearly until it reached a maximum level. In the specification of groundwater costs, we included subsidence costs, which depend on the minimum historic water level. For the pro-

vision of surface water, a spatial element was incorporated so that the transmission costs of surface water depend on the spatial characteristics of the service area. The spatial allocation problem was solved first. We showed that when x percent of the water supply comes from the surface source, the x percent of the population closest to the water treatment plant should be serviced exclusively with surface water while the rest of the city relies on groundwater. How far surface water needs to be delivered depends on spatial parameters of population density, density gradients, and city limits. This result was incorporated in the general social planning dynamic optimization problem. Consumers were assumed to be identical and the social planner was subject to an equity constraint so that all individuals received the same amount of water, regardless of their location.

The theoretical structure of the model and its expected solutions were explored first. As demand grows, the water level of the aquifer should go down and subsidence should occur until the social marginal cost of groundwater reaches the marginal cost of surface water. When this occurs, the two sources are used conjunctively as long as their social marginal costs remain equal. Because of increasing returns to scale in surface water treatment, a period may follow during which water levels increase and subsidence halts. However, decreasing returns to dispersion eventually outweigh returns to scale and, as the marginal cost of surface water increases, groundwater levels decline once again. If new plants cannot be built, subsidence eventually resumes and continues until the aquifer reaches its minimum pressure. After that point, groundwater use must be equal to the recharge rate. The possibility of building a new treatment plant at a different location whenever decreasing returns to dispersions more than outweigh increasing returns to scale would permit long run groundwater levels to stay permanently above their minimum level.

To run the model, all benefits, costs, and transition equations were specified. A chapter was devoted to the specification of demand. Inter-temporal optimization of discounted net

benefits commonly uses pre-estimated demand elasticities to construct a consumer surplus. This approach is not satisfactory when elasticities change endogenously over time. Flexible functional forms, normally used to obtain nonconstant elasticities, are not well suited for out-of-sample forecasts. Moreover, they involve many parameters that render sensitivity analysis too complex. The properties of a parsimonious alternative demand specification, the Stone-Geary, were explored. Its ability to also provide flexible price elasticity was evaluated in an econometric exercise where results from the Stone-Geary were compared to results from a flexible functional form. A panel data on Texas Communities in the early eighties was used in the empirical study where demand was estimated using a random effect panel data estimator. The Stone-Geary was found to provide elasticity estimates that retained their seasonality and were decreasing with price. A water threshold was obtained as a by-product of the Stone-Geary. This threshold corresponds to a level of water use below which demand is not responsive to price changes. Despite its conditional nature, this parameter can prove useful in the design of water conservation policies.

A dynamic simulation of the theoretical model was carried out. Houston being the empirical focus and main motivation for this research, parameter values were specified using extensive information on Houston's population, water demand, income, water levels in wells, historical surface water and groundwater average costs, and subsidence costs estimates from the 1970's. Dynamic programming numerical techniques were used to solve the Bellman equation. We assumed a time invariant state starting in year 2050. The value function was numerically approximated with Spline techniques. The fixed point of the Bellman equation was found using iterative constrained optimization techniques. Backward induction was used to determine the full paths of the control and state variables.

The simulation results confirm the analytical results. All simulations show a decrease in groundwater levels before surface water is used. When water demand is large enough, a

complete switch to surface water occurs. Water levels in the aquifer go up linearly and, soon after it becomes optimal to use some groundwater is conjunction with surface water. Water levels still increase but at a decreasing rate as groundwater use increases but still remains below the recharge rate. Eventually, the use of groundwater increases above recharge and water levels decrease again. Given the time horizon (2050) and the maximum population imposed at the end of the planning period, water levels below the level of subsidence previously reached were never obtained in the simulations. Sensitivity analysis showed that results were quite sensitive to the subsidence costs and the rate of discount but the basic shape of the paths remained the same. The simulation with no subsidence costs showed that subsidence cost alone would not justify any groundwater conservation since the cost of groundwater, inclusive of the pumping externality is always below the cost of surface water. In this case, no surface water is used in this case until the bottom of the aquifer is reached. The sensitivity analysis with respect to spatial variables revealed, contrary to our expectation, that changes in density gradients and city limits, keeping population constant, does not have a significant impact on the results. This fact was explained by the fact that distance to service for surface water also depends on the proportion of surface to groundwater used. When a change in a spatial variable cause the cost of surface water to decrease, more surface water is used relative to groundwater and the service area becomes greater.

The final chapter addresses policy issues using a case study of the Houston-Galveston region. The history of groundwater and surface water use as well as the groundwater regulation plans of the Houston-Galveston Coastal Subsidence District were reviewed and discussed in the light of our previous findings. We compared observed data to simulated paths of groundwater use, surface water use, and resulting water levels in the aquifer from 1950 to 1995. Houston started using surface water earlier than the model predicts and the level of subsidence reached before water level started rising again is lower than in most of the sim-

ulations. Levels of subsidence similar to the observed are obtained with a low discount rate or if subsidence costs are very large. As the model does not incorporate spatial differences in subsidence it is difficult to make definite statements about whether Houston controlled subsidence optimally. Yet, it is clear from the simulations that Houston did a enough to address the subsidence problem and there should be no concern when water levels start downward again.

We argued that positive feedback from districts concerned with increasing costs provided incentives to revise the subsidence plan in 1992 in a manner more consistent with a cost-minimizing path, taking into account potential increasing returns in surface water treatment. It is also argued that the policy to allocate a fair share of the safe groundwater yield to each area is not optimal because it does not take into account the differences in costs of providing water from alternate sources. A system of tradable permits will increase efficiency without distributional consequences if there is enough hydrologic connection between the zones. Given the estimated transmissivities of the Evangeline aquifer, the permit system would have to take into consideration that it may take several years for water to travel between two areas at opposite side of the region. Further research is needed to assess the practicality and potential effectiveness of such groundwater permits in the area.

REFERENCES

[1] Amigues, J.P., G. Gaudet; and M. Moreaux (1995). On Some Specificities of the Valuation of Semi-Renewable Natural Resources. (in French - English abstract). *Annales d'Economie et de Statistiques*. Issue 39: pp. 107-21.

[2] Arrow, K. J. and A. C. Fisher (1974). Environmental Preservation, Uncertainty and Irreversibility. *Quarterly Journal of Economics* Vol. 87: pp. 312-19.

[3] Baltagi, Badi and Young-Jae Chang (1994) Incomplete Panels: A Comparative Study of Alternative Estimators for the Unbalanced One-Way Error Component Regression Model. *Journal of Econometrics*, 62(2): 67-89.

[4] Baltagi, Badi, (1995) *Econometric Analysis of Panel Data*, New York: John Wiley & Sons, Inc.

[5] Blackorby, C. and D. Donaldson (1985). Consumers' Surplus and Consistent Cost Benefit Tests. *Social Choice and Welfare* Vol. 1: pp. 251-62.

[6] Brown, Gardner Jr. and Robert Deacon (1972). Economic Optimization of a Single-Cell Aquifer. *Water Resources Research*, Vol. 8(3): pp. 557-564.

[7] Brown, Bryan W. and Mary Beth Walker (1989) "The Random Utility Hypothesis and Inference in Demand Systems," *Econometrica*, Vol. 57(4): pp. 815-829.

[8] Buras, Nathan (1972). *Scientific Allocation of Water Resources*. New York: American Elsevier.

[9] Burt, Oscar R. (1964). The Economics of Conjunctive Use of Ground and Surface Water. *Hilgardia* Vol. 36(2): pp. 31-111.

[10] Burt, Oscar R. and Ronald G. Cummings (1970). Production and Investment in Natural Resource Industry. *American Economic Review* Vol. 60 (4): pp. 577-90.

[11] Buse, A. (1973) Goodness of fit in Generalized Least Square Estimations. *American Statistician*, 27: 106-108.

[12] Carr, Jerry E., Walter R. Meyer, William M. Sandeen, and Ivy R. McLane: US Geological Survey, (1985). *Digital Models for Simulation of Ground-Water Hydrology of the Chicot and Evangeline Aquifers along the Gulf Coast of Texas*. Texas Department of Water Resources, Report No. 289.

[13] Chung, Jae Wan, (1994) *Utility and Production Functions: Theory and Applications*, Cambridge: Blackwell Publishers.

[14] Clark, Robert M. and Richard Stevie (1981). A Water Supply Cost Model Incorporating Spatial Variables. *Land Economics*, Vol. 57, No. 1: pp. 18-32.

[15] Clarke, H. R. and R. M. Shrestha (1987). Renewable Resource Management with a Backstop Substitute and Nonautonomous Prices. *Journal of Environmental Economics and Management* Vol. 14: pp. 159-182.

[16] Cochran, Richard and Arthur W. Cotton, (1985). Municipal Water Demand Study, Oklahoma City and Tulsa, Oklahoma. *Water Resources Research*, 21(7): 941-943.

[17] Conrad Jon M. and Colin W. Clark. (1987). *Natural Resource Economics: Notes and Problems* Problems.. Cambridge University Press: Cambridge.

[18] Crocker, Thomas D.(1966). The Structuring of Atmospheric Pollution Control Systems. In *The Economics of Air Pollution*, Harold Wolozin, Ed. Norton: New York.

[19] Custodio E. and A Gurgui eds (1989). Groundwater Economics: Selected Papers from a United Nations Symposium held in Barcelona, Spain. *Development in Water Science* (39). Elsevier: Amsterdam.

[20] Dajani, J.S. and R.S. Gemmel (1973). Economic Guidelines for public Utilities Planning. *Journal of the Urban Planning and Development Division* No. 99 (Sept.): pp. 171-82.

[21] Dales, John H. (1968). *Pollution, Property, and Prices*. University of Toronto Press: Toronto.

[22] Danielson, Leon E., (1979). An Analysis of Residential Demand For Water Using Micro Time-Series Data. *Water Resources Research*, 15(4): 763-767.

[23] Dasgupta, P. S. and G. M. Heal (1979). *Economic Theory and Exhaustible Resources*. Cambridge University Press: Cambridge (UK).

[24] Dasgupta, P. S. and G. M. Heal (1982). *The Control of Resources*. Harvard University Press: Cambridge.

[25] Dasgupta P.S. and J. Stiglitz (1981) Resource Depletion Under Technological Uncertainty. *Econometrica* 49(1): pp. 85-104.

[26] Deaton, Angus and John Muellbauer, (1980) *Economics and Consumer Behavior*, Cambridge University Press: Cambridge (UK); New York.

[27] Deaton, Angus and John Muellbauer, (1980b) An Almost Ideal Demand System. *American Economic Review*, 70: 312-326.

[28] DeLucia, R., E. McBean, and J. Harrington (1978). A water Planning Model With Multiple Time Pollutant and Source Capabilities. *Water Resources Research*, 14(1): pp.9-14.

[29] Diewert,W. E., (1973) Separability and a Generalization of the Cobb-Douglas Cost, Production, and Indirect Utility Functions. *Technical Report* 86, IMSSS, Stanford University.

[30] Diewert, W. E. and T. J. Wales (1987). Flexible Functional Forms and Global Curvature Conditions. *Econometrica*, 55(1): 43-68.

[31] Diewert, W. E. and T. J. Wales (1988). A Normalized Quadratic Semiflexible Functional Form. *Journal of Econometrics*, 37(3): 327-342.

[32] Edmonton, Barry (1975) *Population Distribution in American Cities*. Heath and Company: Lexington.

[33] Ford, J. L. and J.J. Warford (1969). Cost Functions for the Water Industry. *Journal of Industrial Economics*, 18(1): 53-63.

[34] Foster, Henry S. and Bruce R. Beattie, (1979) Urban Residential Demand for Water in the United States. *Land Economics*, 55(1): 43-58.

[35] Gabrysch, Robert K. (1969) Land Surface Subsidence in the Houston-Galveston Region, Texas. *Proceedings of the International Symposium on Land Subsidence*. Tokyo.

[36] Gao, Weiyu, Peter R. Hartley, and Robin C. Sickles (1999). Optimal Dynamic Production Policy: The Case of a Large Oil Field in Saudi Arabia. Mimeo, Rice University.

[37] Gibbs, Kenneth C. (1978). Price Variables in Residential Water Demand Models. *Water Resources Research* Vol. 14(1): pp. 15-18.

[38] Gisser, Micha and David A. Sanchez (1980). Competition Versus Optimal Control in Groundwater Pumping. *Water Resources Research*. Vol. 16 (4): pp. 638-42.

[39] Green, William H., (1997) *Econometric Analysis: Third Edition*, Prentice-Hall, Inc.: New Jersey

[40] Griffin, Ronald C. and Chan Chang. (1989). Community Water Demand in Texas. *Technical Reports* #149, Texas Water Resources Institute: Texas A&M University.

[41] Griffin, Ronald C. and Chan Chang, (1990). Pretest Analyses of Water Demand in Thirty Communities. *Water Resources Research*, 26(10): 2251-2255.

[42] Griffin, Ronald C. and Chan Chang, (1991). Seasonality in Community Water Demand. *Western Journal of Agricultural Economics*, 16(2): 207-217.

[43] Guilkey, David K., Lovell,, C. A. Knox, and Robin C. Sickles, (1983). A Comparison of the Performance of Three Flexible Functional Forms. *International Economic Review*, 24(3): 591-616.

[44] Harris-Galveston Coastal Subsidence District (1983). *Subsidence '83*. HGCSD Publication: Friendswood, Texas.

[45] Harris-Galveston Coastal Subsidence District (1985). *District Plan, Adopted November 1985*. HGCSD Publication: Friendswood, Texas.

[46] Harris-Galveston Coastal Subsidence District (1994). *Ground Water Report, Year Ending December 31, 1993. Adopted May 11, 1994*. HGCSD Publication: Friendswood, Texas.

[47] Hartley, Peter R. (1996). Value Function Approximation in the Presence of Uncertainty and Inequality Constraints: An Application to the Demand for Credit Cards. *Journal Of Economic Dynamics and Control.* Vol. 20: pp. 63-92.

[48] Hartley, Peter R. and Albert S. Kyle (1989). Equilibrium Investment in an Industry with Moderate Investment Economies of Scale. *The Economic Journal*, Vol. 99: pp. 392-407.

[49] Hausman, Jerry A. (1981). Exact Consumer Surplus and Dead Weight Loss. *American Economic Review* Vol. 71 (4): pp. 662-76.

[50] Hausman, Jerry A. and Whitney K. Newey (1995). Nonparametric Estimation of Exact Consumer Surplus and Deadweight Loss. *Econometrica*, 63(6): 1445-1476.

[51] Heal, J. M. (1976). The relationship between Price and Extraction Cost for a Resource with a Backstop Technology. *Bell Journal of Economics* Vol. 7: pp. 371-8.

[52] Henry C. (1974). Investment Decisions Under Uncertainty: The Irreversibility Effect. *American Economic Review*, Vol. 64: pp. 1006-12.

[53] Hotelling H. (1931). The Economics of Exhaustible Resources. *Journal of Political Economy* Vol. 39: pp. 137-75.

[54] Howe, C. W., and F. P. Linaweaver, Jr., (1973). The Impact of Price on Residential Water Demand and its Relation to System Design and Price Structure. *Water Resources Research*, 3(1): 13-32.

[55] Johnson, Marvin B. (1979). Community Income, Intergovernmental Grants, and Local School District Fiscal Behavior. In *Fiscal Federalism and Grants-in-Aid*, Mieszkowski, Peter and William H. Oakland, Eds., pp.51-77.

[56] Jones, Lonnie L. and James Larson (1975). Economic Effects of Land Subsidence Due to Excessive Groundwater Withdrawal in the Texas Gulf Coast Area. Texas Water Resources Institute. *Technical Report* No. 67. Texas A&M University: College Station, Texas.

[57] Kemp, M. C. and N. Van Long (1980). On Two Folk Theorems Concerning the Extraction of Exhaustible Resources. *Econometrica* Vol. 48(3): pp. 663-73.

[58] Knapp, K and Lars Olson (1995). The Economics of Conjunctive Groundwater Management with Stochastic Surface Supplies. *Journal of Environmental Economics and Management* Vol. 28: pp. 340-56.

[59] Lewis, T. R. (1982). Sufficient Conditions for Extracting Least Cost Resource First. *Econometrica* Vol. 50(4): 1081-1083.

[60] Manne, Alan S. (1967) *Investments for capacity expansion; size, location, and time-phasing*, MIT Press: Cambridge.

[61] Martin, Randolph C. and Ronald P. Wilder (1992). Residential Demand for Water and the Pricing of Municipal Water Services. *Public Finance Quarterly*, Vol. 20(1): pp. 93-102.

[62] Martin, William E. and John F. Thomas (1986). Policy relevance in Studies of Urban Residential Water Demand. *Water Resources Research*, Vol. 22 (13): pp. 1735-41.

[63] McGuire, Martin C. (1979). The Analysis of Federal Grants into Price and Income Components. In *Fiscal Federalism and Grants-in-Aid*, Mieszkowski, Peter and William H. Oakland, Eds. , pp.31-50.

[64] Mieszkowski, Peter and Barton Smith (1991) Analyzing Decentralization: The Case of Houston. *Regional Science and Urban Economics*, Vol. 21(2): pp. 183-199.

[65] Mieszkowski, Peter and Edwin S. Mills (1993) The Causes of Metropolitan Suburbanization. *Journal of Economic Perspectives*, Vol. 7(3): pp. 135-147.

[66] Mills, Edwin S. (1972) *Studies in the Structure of The Urban Economy*. Resources for the Future/John Hopkins University Press: Baltimore

[67] Montgomery, W. David (1972). Markets in Licences and Efficient Pollution Control Programs. *Journal of Economic Theory*, pp. 395-418.

[68] NHCWSC (1987). *North Harris County Water Supply Corporation: Surface Water Conversion Plan*. Report prepared by Pate Engineers/Jones&Carter, Joint Venture, submitted to NHCWSC, June 1987.

[69] Nickum, James E. and William Easter (1994) The Maturing Metropolitan Water Economies. In *Metropolitan Water Use Conflicts in Asia and the Pacific*. Nickum, James E. and William Easter eds., Westview Press: Boulder

[70] Powell, A., (1974) *Empirical Analytics of Demand Systems*, Lexington, MA: D. C. Heath.

[71] Renzetti, Steven. (1992). Evaluating the Welfare Effects of Reforming Municipal Water Prices. *Journal of Environmental Economics and Management*, Vol. 22: pp. 147-63.

[72] Rubio, Santiago, Constantino Martinez, and Juan Castro (1994). Optimal Management of Groundwater with Increasing Demand. *Revista Española de Economia*. Special Issue 1994: pp.161-89.

[73] SAS ETS Software(1996) *Changes and Enhancements for Release 6.12*. SAS Institute Inc.

[74] Stavins, Robert N. (1998). What Can We Learn from the Grand Policy Experiment: Lessons from SO2 Allowance Trading. *Journal of Economic Perspectives*, Vol. 12(3): pp.. 69-88.

[75] Stokey, Nancy L. and Robert E. Lucas, Jr., with Edward C. Prescott (1989). *Recursive methods in economic dynamics*. Harvard University Press: Cambridge.

[76] Stone, R. (1954). *The Measurement of Consumers' Expenditure and Behavior in the United Kingdom, 1920-1938*, Volume I, Cambridge University Press: Cambridge (UK).

[77] Terrell, Dek (1996). Incorporating Monotonicity and Concavity Conditions in Flexible Functional Forms. *Journal of Applied Econometrics*, 11(2): 179-194.

[78] Teutsch, John S. (1978). *Subsidence in the Houston-Galveston Region: A Comprehensive Analysis*. M. A. Thesis, Rice University: Houston, Texas.

[79] Texas Water Development Board (1990). *Water for Texas*. TWDB, pp.2-8.

[80] Thomas, J. F. and W. E. Martin (1989). Mining of Aquifers near Metropolitan areas: Toward a General Framework for Policy Analysis. In *Groundwater Economics*, E. Custodio and A Gurgui Eds: pp. 397-408. (see Custodio)

[81] Tsur, Y and A. Issar (1989). The Buffer Role of Ground Water When Supply of Surface Water is Uncertain. in *Groundwater Economics*, E. Custodio and A Gurgui eds: pp. 373-80. (see Custodio)

[82] Tsur, Y and Amos Zemel (1995). Uncertainty and Irreversibility in Groundwater Resource Management. *Journal of Environmental Economics and Management* Vol. 29: pp. 149-61.

[83] Warren, John, Lonnie Jones, Ronald Lacewell, and Wade Griffin (1975). External Costs of Land Subsidence in the Houston Bayton Area. *American Journal of Agricultural Economics*, Vol. 57(3): pp.450-455.

[84] Waltham, A. C. (1984).) *Ground Subsidence*, New York: Chapman and Hall.

[85] Wansbeek, Tom and Arie Kapteyn (1989) Estimation of the Error-Components Model with Incomplete Panels. *Journal of Econometrics*, 41: 341-361.

[86] West S. W. and W. L. Broadhurst (1983). Rio Grande Region. In *Groundwater Resources of the United States*, Todd, David K. ed. Premier Press: Berkeley.

[87] Willig, R. (1976). Consumer Surplus without Apology. *American Economic Review* Vol. 66: pp. 589-97.

[88] World Bank, 1993. Water Resources Management. *World Bank Policy Paper*. Washington D.C.

[89] World Development Report (1992). *Development and the Environment*. New York: Oxford University Press

APPENDIX

-Figure A1 - Houston-Galveston Area Aquifers: Hydrogeologic Cross–Section -Figure A2 - Surface water supply for the City of Houston.

-Figure A3 - Subsidence Estimates: 1987-1995

-Figure A4 - Subsidence Estimates: 1973-1995

-Figure A5 - Subsidence Estimates: 1906-1995

-Figure A6 - Zones of the 1985 Houston Galveston Subsidence District Plan

-Figure A7 - Zones of the 1992 Houston Galveston Subsidence District Plan

Figure A1: Houston-Galveston Area Aquifers: Hydrogeologic Cross Section

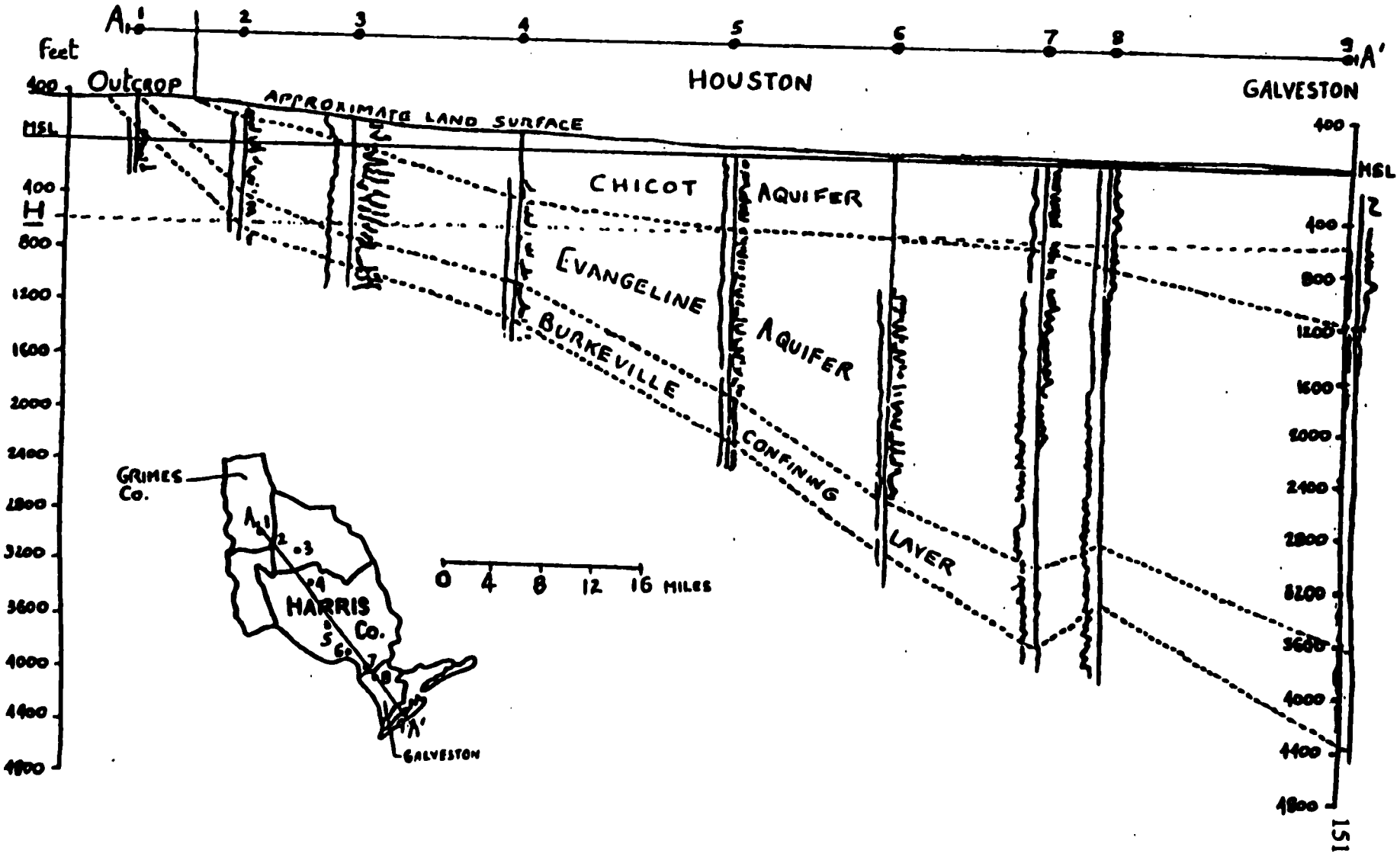


Figure A2: Surface water supply for the City of Houston

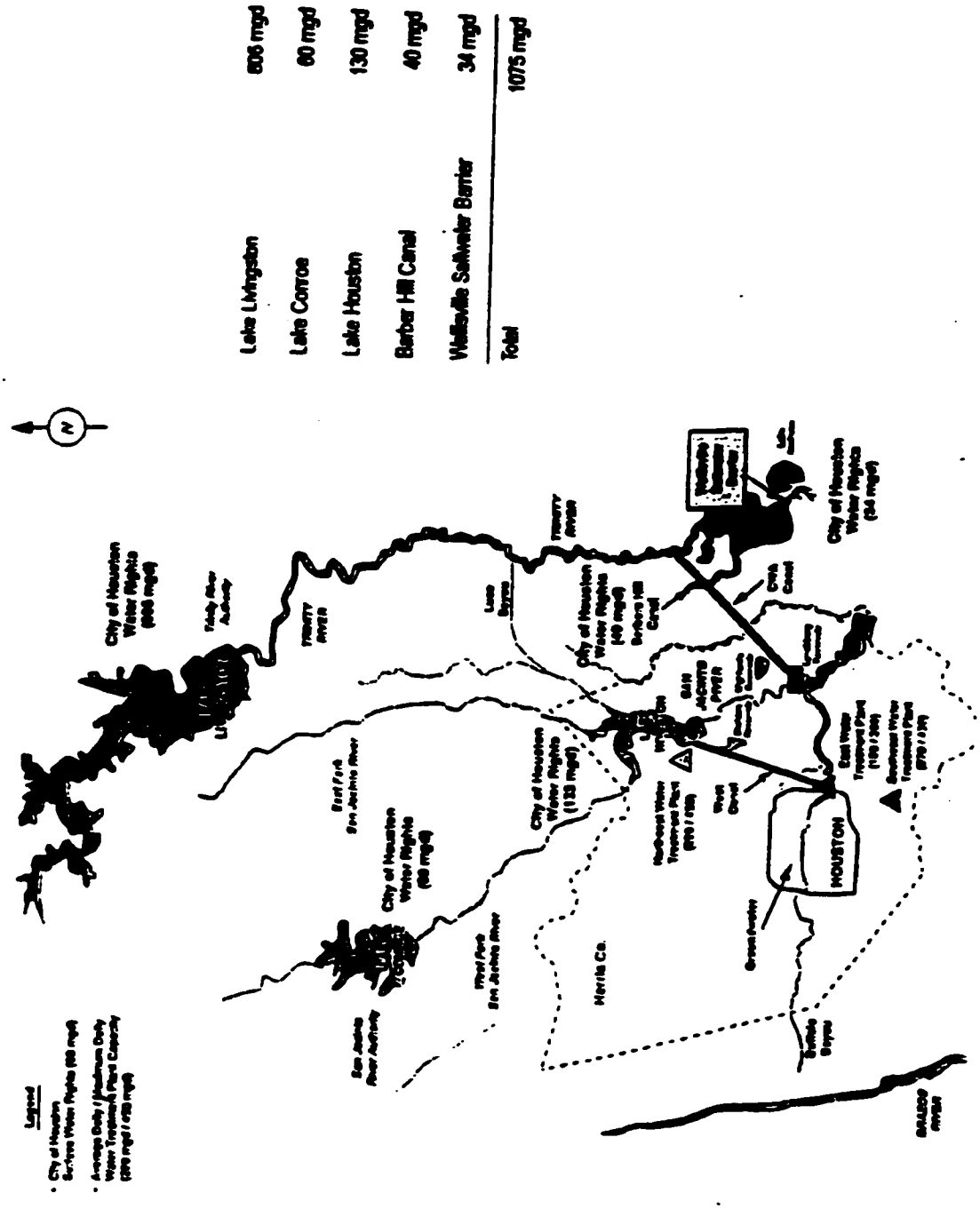


Figure A3: Subsidence Estimates: 1987-1995

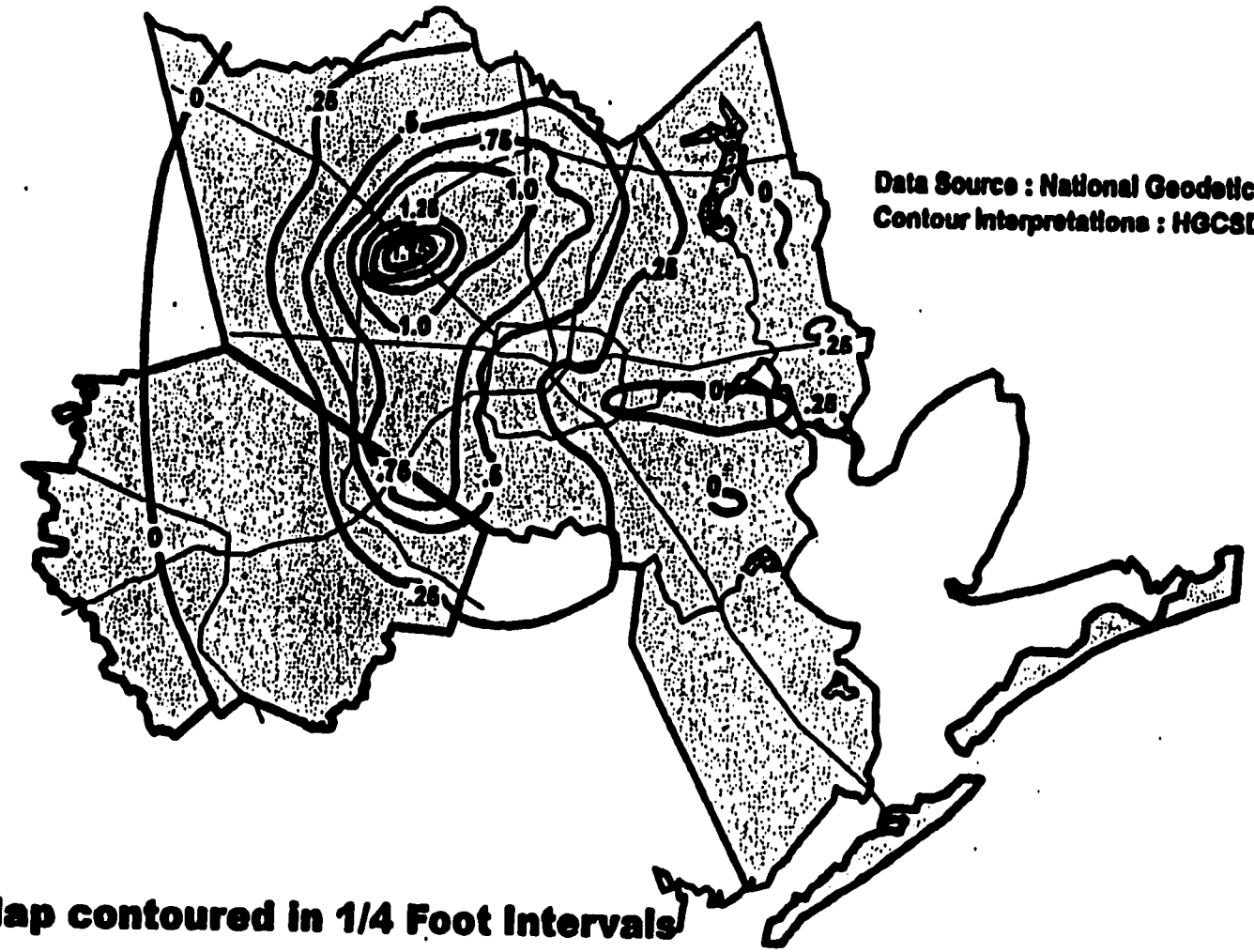


Figure A4: Subsidence Estimates: 1973-1995

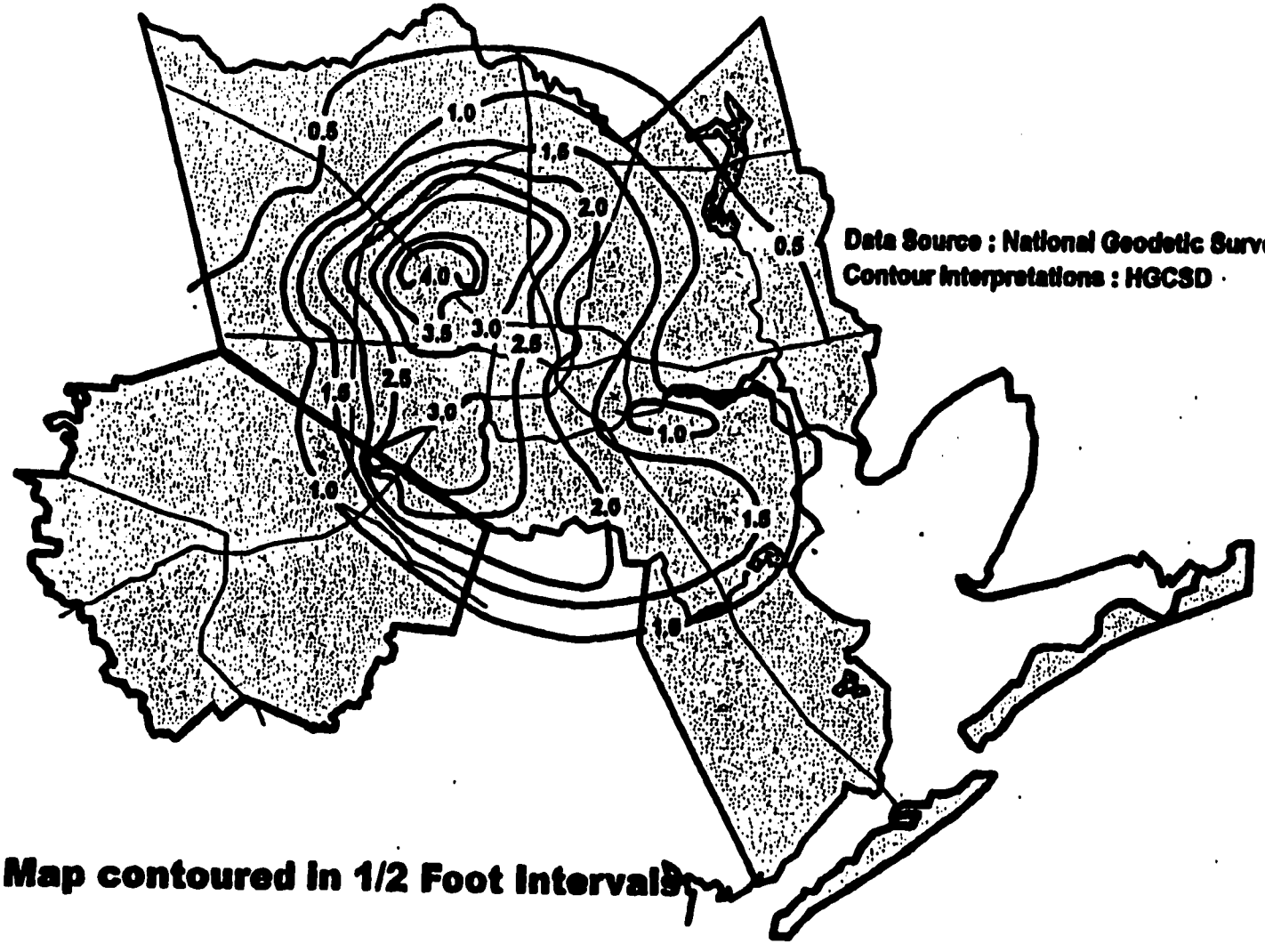


Figure A5: Subsidence Estimates: 1906-1995

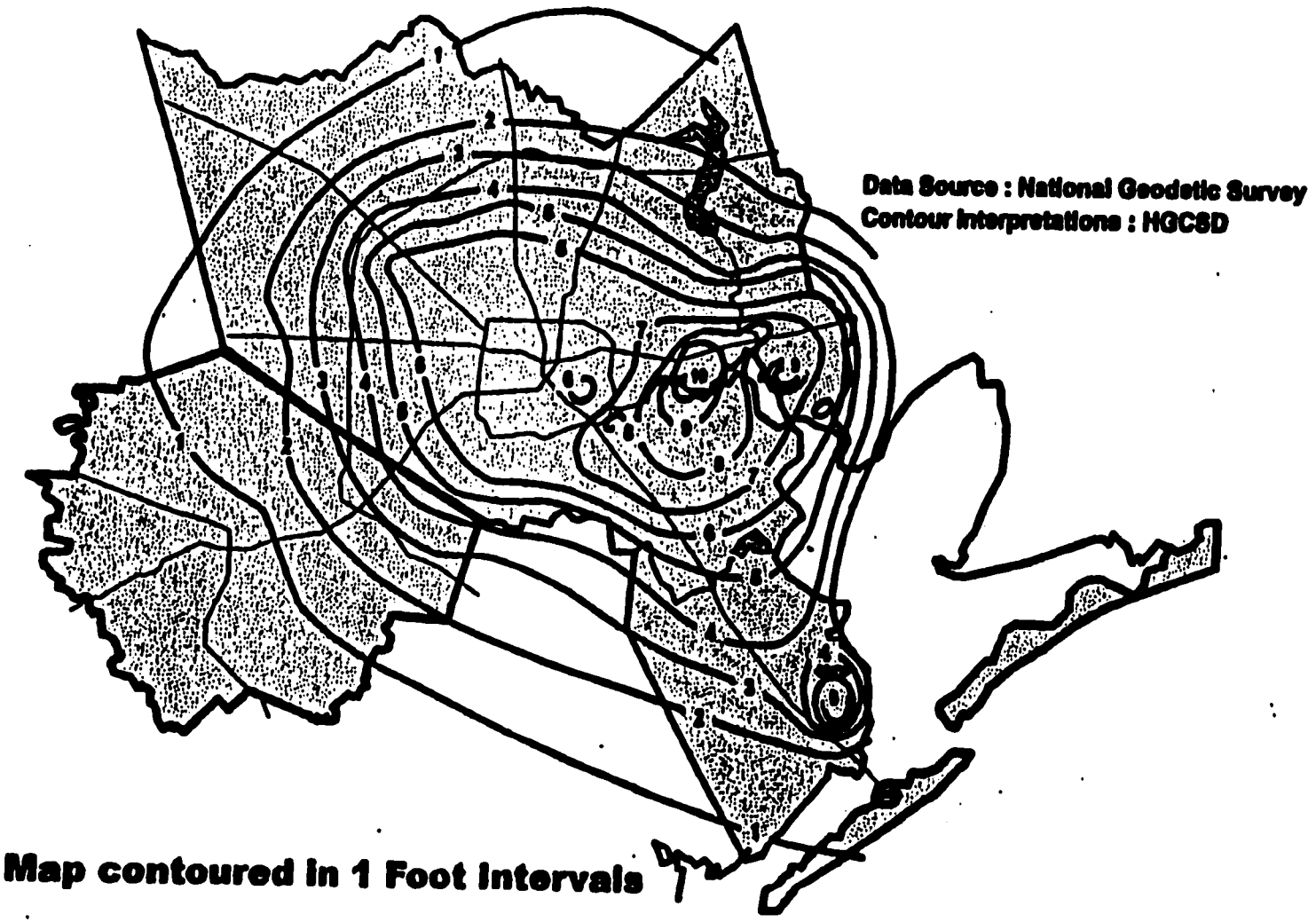
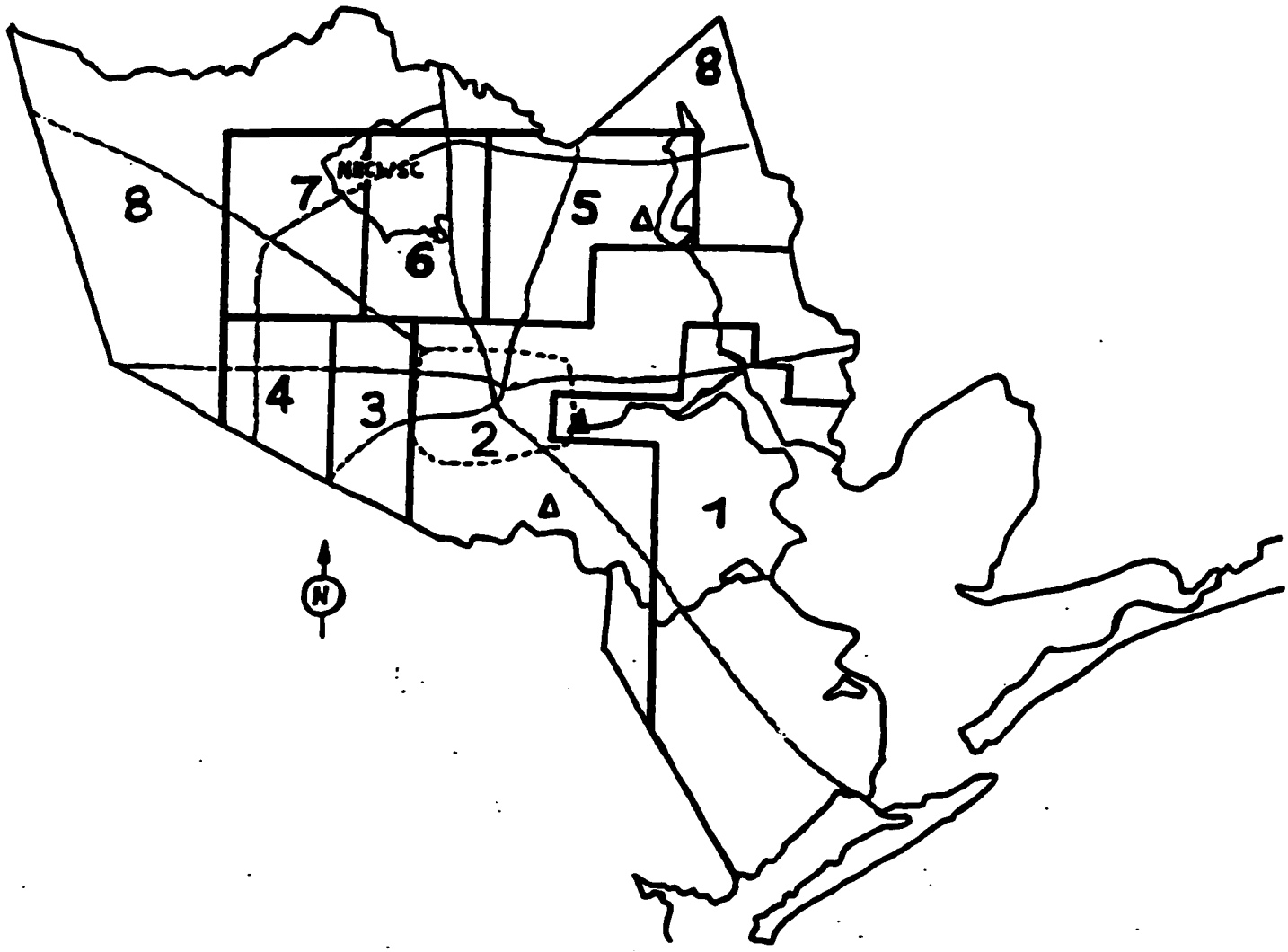


Figure A6: Zones of the 1985 Houston Galveston Subsidence District Plan

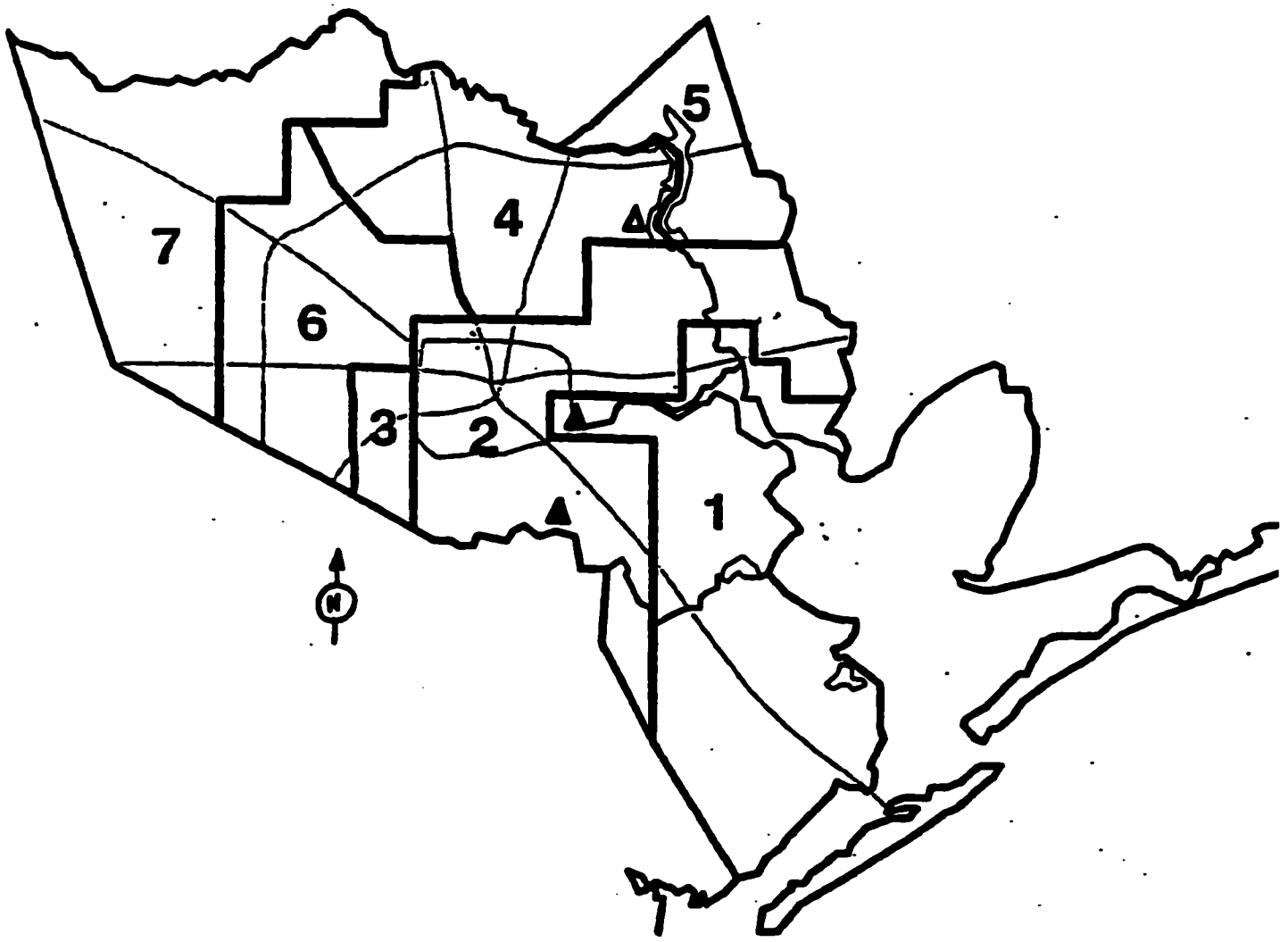


▲ Location of existing treatment plants

Δ Location of future treatment plants

Sources: HGCSD (1885), NHCWSC (1987), City of Houston.

Figure A7: Zones of the 1992 Houston Galveston Subsidence District Plan



▲ Location of existing treatment plants

△ Location of future treatment plants

Sources: HGCSD (1994), City of Houston