An Evaluation of the Effects of a Pricing Policy on the Water Consumption of Heterogeneous Households in Seattle

Yuin-Jen David Hsu

A dissertation submitted in partial fulfillment
of the requirements for the degree of

Doctor of Philosophy

University of Washington

2010

Program Authorized to Offer Degree: Urban Design & Planning Group

University of Washington
Graduate School

This is to certify that I have examined this copy of a doctoral dissertation by

Yuin-Jen David Hsu

and have found that it is complete and satisfactory in all respects,
and that any and all revisions required by the final
examining committee have been made.

Chair of the Supervisory Committee:

Paul A. Waddell

Reading Committee:

David F. Layton

Paul A. Waddell

Richard O. Zerbe

Date:

University of Washington

Abstract

An Evaluation of the Effects of a Pricing Policy on the Water Consumption of Heterogeneous Households in Seattle

Yuin-Jen David Hsu

Chair of the Supervisory Committee:
Professor Paul A. Waddell
Urban Design & Planning

In 2001, the City of Seattle implemented a new pricing policy for residential water consumption, intended to target high water users. This dissertation evaluates the impact of this policy and contributes to the water demand literature in three ways. First, this dissertation extends the use of instrumental variables to represent the effect of nonlinear price structures on heterogeneous users. Correlated random coefficient models, specified with coefficients that vary by groups such as households and neighborhoods, allow a realistic specification of heterogeneity while addressing simultaneity bias. Price elasticity estimates for the entire population range from -0.15 to -0.52, depending on the model used and the point of consumption, and are consistent with the literature. Second, this dissertation uses a hierarchical linear model to link the varying intercepts and coefficients to the underlying properties of heterogeneous individuals and groups, and to allow meaningful, policy-relevant interpretation of the estimates, as well as a sensitivity analysis of factors affecting the policy impacts. A household-level model finds an average price elasticity of approximately -0.41; the average household price elasticity also varies as expected with lot size and house value. Third, based on these predictive model results, this dissertation explores the statistical properties of welfare measures for heterogeneous groups in order to characterize the distributive impact of this pricing policy. Estimated losses in consumer surplus as a result of the pricing policy range from 5-45% of existing total bills.

TABLE OF CONTENTS

LIST OF FIGURES

Figure Number Page

LIST OF TABLES

Table Number Page

GLOSSARY

2SLS: Two-Stage Least Squares

AIC: Akaike Information Criterion

ANOVA: Analysis of Variance

CCF: 100 cubic feet of water, equivalent to 748 gallons

CRC: Correlated Random Coefficients

CS: Consumer's Surplus

CV: Compensating Variation

DCC: Discrete-Continuous Choice

DIC: Deviance Information Criterion

DV: Difference Variable

EV: Equivalent Variation

HLM: Hierarchical Linear Modeling

IV: Instrumental Variable

LATE: Local Average Treatment Effect

MP: Marginal Price

NOAA: National Oceanographic & Atmospheric Administration

OLS: Ordinary Least Squares

PDF: probability density function

PMF: probability mass function

RMSE: Root Mean Square Error

RE: Random Effect

SPU: Seattle Public Utilities, the municipally-owned utility for drinking water, wastewater, and solid waste in the City of Seattle.

WWIV: Wilder and Willenborg (1975) Instrumental Variable

ACKNOWLEDGMENTS

I'd like to express my sincere appreciation and gratitude to my dissertation committee. To my advisor, Paul Waddell, I am grateful for his superb and subtle guidance in my research and career. To the rest of my dissertation committee – David Layton, Richard Zerbe, and John Findlay – I am grateful for their generous sharing of ideas, time, and expertise. I hope to do as well in my research and teaching in the future.

Particular thanks go to Bruce Flory of Seattle Public Utilities. Without his intellectual curiosity and openness, this project would not have been possible.

I feel very fortunate to have been part of several communities of scholars and friends in my time at the University of Washington. Being part of these communities made the great experiences in my education, research, work, and life better and possible in every way. Things I remember and remain grateful for, include but are not limited to: ambitious and challenging teaching; working together on interesting projects; discussing research; numerous software and research tips; or just shared good humor and company over innumerable seminars, practice talks, walks, lunches, cups of tea, and snacks:

- In the UrbanSim lab, thanks to Alan Borning, Jin Hyun Hong, Travis Kriplean, Brian Lee, Hana Sevcikova, and Liming Wang, for creating an excellent and fun environment for thinking, working, and socializing.

- In the Interdisciplinary PhD program, thanks to Ahmed Al-Noubani, Alon Bassok, Andrew Bjorn, Emad Dawwas, Sarah Dooling, Joel Franklin, Adrienne Greve, Phil Hurvitz, Yan Jiang, Junfeng Jiao, Lin Lin, Karis Puruncajas, and Danielle Spirandelli, for great discussions, guidance, and camaraderie. Special thanks to Bob Mugerauer for his unique combination of practical and philosophical insight and great enthusiasm. Jean Rogers, Jan Brooks, and Krista Bargsten provided friendly help everyday. Anne

Vernez Moudon can always be counted on to tell it like it is. Thanks to Marina Alberti for bringing me to Seattle in the first place.

- Thanks to Kristin Broms, Hans Nesse, Gail Potter, and Kevin See for welcoming me into the company of statistics boot camp; in retrospect, it was kind of fun.

- In the CSSS program and Statistics, thanks to Adrian Dobra, Mark Handcock, Darryl Holman, Michael Perlman, Adrian Raftery, and Thomas Richardson for terrific teaching and opening up still more things and fields to study.

- For good company and help far afield from urban planning, thanks to Wahid Abdallah, Anne Buffardi, Colleen Chrisinger, Paul Corcoro, Brian Halbert, Hsiao-Chi Hsu, Alex Kirk, Katie Lucia, and Siobhan Mattison.

- In TRAC, thanks to Mark Hallenbeck and Mike Richards for generously hosting me for the last, critical year of graduate school.

Also, special thanks to Robert Balzar, Steve Lush, Michael Little, and Glenn Atwood, for giving me great responsibilities and opportunities at Seattle City Light. Working downtown was a welcome relief from graduate school, and exposed me to the really exciting things happening in energy and conservation. I can say that the experience has really shaped my future career path.

Finally, none of my education and achievements would would have been possible without the support and encouragement of my family. My mother, father, uncle, and grandfather are the first four professors and doctors in our family, and I feel honored to follow on this well-trod path as the fifth. Thank you to my sister Jennifer and my brother-in-law John, and their family, for encouraging me to pursue what I'm passionate about. Also, thanks to my parents-in-law, Lynne and John, who have also been unfailingly supportive and complimentary of my studies.

Lastly, it is typical to honor the author's immediate family for their support on this project. Usually it sounds as if the author has heroically wrestled alone with great ideas while the family provided tea, sympathy, or comic relief; or else that the family has operated as a Marxist/Borg unit collective, dedicated entirely to meeting some conceptual production quota determined by a far-off committee. I hope that we've avoided these extremes, because as much as doctoral education has changed me over the past five years, I am also certain that gaining a family has changed me even more. And although this dissertation reflects my particular academic interests, I also hope that my work has and will continue to enable us to appreciate our everyday lives together, and to pursue our individual aspirations and collective dreams. This dissertation would also simply mean less without my family's help, support, and love. So, I am grateful to our family for the best of company over the past few years, and look forward to our shared future. I offer my profound thanks, to and for, my smartful and amazing wife Katherine, our beautiful and funny daughter Beatrice, and our rambunctious and under-appreciated dog Penny.

Any errors in this document, of course, are mine alone. I look forward to a long academic career of correcting them and making new ones.

DEDICATION

To our family – which grows everyday – Katherine, Beatrice, and Penny

Chapter 1

INTRODUCTION

In 2001, Seattle Public Utilities (SPU), the municipally-owned water utility for the City of Seattle, changed its pricing policies for water by adding a new and substantially higher price block to their existing block rate price structure. This new price only affected the highest consumers of water, and was often referred to as a 'shock rate' intended to signal forcefully to particular users their relatively high water consumption. The research question that this dissertation seeks to answer is: how did this change in pricing policy affect water consumption for different users?

Answering this question contributes to the academic literature in several fields, including environmental economics, statistics, and urban planning. Rigorous analysis of this policy also has broad implications for how we think about managing the use of critical environmental resources in cities – such as water – in a world in which urbanization and climate change are still accelerating. This dissertation develops analytical tools to aid in the prediction of household water demand, to determine influences on water demand, and to calculate distributive effects on heterogeneous users, all in order to contribute to the academic literature and wider environmental and policy discussions.

This chapter will 'unpack' this research question into its various parts, in order to connect this research to the existing academic literature, and to demonstrate the relevance of the analytical tools developed in the later research chapters. In the first section below, I will describe the policy intervention and its intended effects at greater length. In the second section following, I will summarize key academic issues related to this policy change as an early preview to the deeper and more technical literature review in Chapter 2. In the third section below, I will summarize the intended, original contributions of this dissertation to the literature from the research presented in Chapters 4-6. In the fourth and final section of this introduction, I will explain how this research is intended to influence future work, by describing the relevance of this research to wider intellectual, environmental, and technological trends. In sum, this chapter introduces the main themes that the subsequent dissertation will explore in much greater depth.

1.1 Policy Description

SPU, like many public utilities, uses a price schedule in which different prices are charged according to the season, classes of use, and the total water consumption. Different seasonal prices are charged because of widely varying seasonal demands for water, which is lower in the winter and higher in the summer. Different prices are also charged to commercial and industrial users versus residential users, although this dissertation will only focus on the residential sector.

The specific policy that this dissertation focuses on was a change in the price schedule for water in the peak summer season for residential users, in which different prices are charged to users depending on their total monthly water consumption. Before 2001, the price structure had two prices that were not greatly different: below 5 ccf per month, SPU charged \$2.48 per ccf of water, and after 5 ccf per month, SPU charged \$3.28 per ccf of water over that threshold.[1] On the whole, water rates have remained quite low over the past 20 years: the typical bimonthly bill for an average user in the summer in this period has been approximately \$25.

In 2001, halfway through the summer, Seattle Public Utilities implemented a third price block, which they occasionally refer to as 'the shock rate'. For users who consumed more than 15 ccf per month, this raised the marginal price of water to \$13.07 per ccf, or almost a 400% increase over the previous marginal price. The intention of this 'shock rate' was to increase the price of water rapidly to get the attention of high water users – in particular those consuming that much unwittingly so – and to give them a significant incentive to change their behavior. Figure 1.1 below shows the prices by block after the price change.

There were good reasons to target specifically high water users. By looking at the

[1] Measured in 2007 real dollars. A ccf is 100 cubic feet or 748 gallons of water.

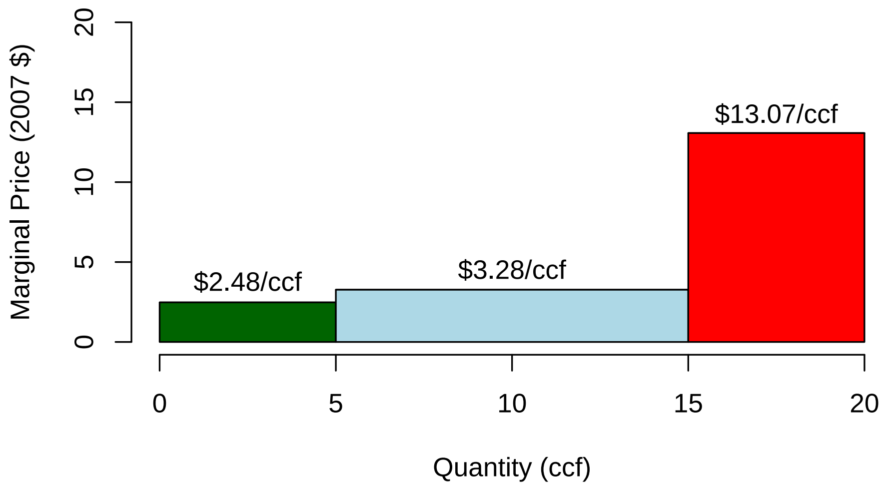

Figure 1.1: SPU Marginal Prices for Water in 2001, adjusted for inflation using CPI 2007 = 1. The green and blue blocks on the left indicate the marginal price for water in the first two blocks, both before and after 2001. The red block on the right represents the new 'shock rate' imposed in 2001 on consumption over 15 ccf per month (Flory, 2008). The threshold for the shock rate was subsequently changed to 18 ccf in later years. Before 2001 the marginal price above this threshold was the same as the middle block, $3.28 per ccf.

distribution of total water consumption by tier in Figure 1.2, we can see that the highest water users consume nearly double their share of water. These distributions show that the highest water consumers represent only 7-15% of the customers, while the distribution shows that they account for almost 20-40% of total water consumption. There are also good reasons to believe that the shock rate had a significant effect on total water consumption. For a typical household in Seattle in the ten years before the price change, average water consumption ranged from approximately 7.1 to 8.8 ccf per month per household. In the six years after the policy was implemented, water consumption steadily dropped from 7.1 ccf per month to 5.7 ccf per month per household in 2007.

The new pricing policy was imposed over a relatively steady background of water consumption and conservation in the City of Seattle. Over the past 20 years SPU has consis-

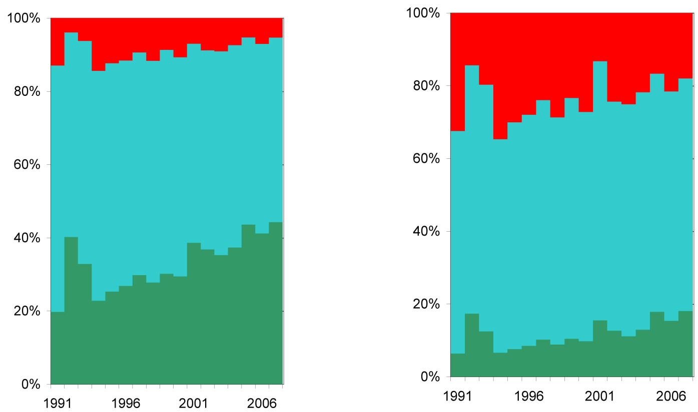

Figure 1.2: Distribution of Water Consumption by Tiers from 1991-2007. The left hand side distribution shows the number of customers by tier, and the right hand side distribution shows the quantities consumed by customers observed in each tier. The left hand side distribution shows that the highest water consumers (in red) represent only 7-15% of the customers, while the right hand side distribution shows that they account for 20-40% of total water consumption. Annual water use was calculated directly from billing data as described in Chapter 3.

tently pursued a combination of water conservation activities, including public information campaigns, education, appliance retrofitting incentives and campaigns, and changes to the building code over the past 20 years. Major droughts in 1992 and 2001 resulted in the declaration of mandatory and voluntary water use restrictions, respectively, but these did not seem to have long-lasting effects afterwards, since water use quickly bounced back in the following years.

Interestingly, many of these activities that SPU and other utilities, such as Seattle City Light, have undertaken mirror different academic theories of the barriers to resource conservation (see, for example, Blumstein et al., 1980, for theories of barriers to energy conservation). We will therefore now examine the key issues in this pricing change, in order to understand why pricing policies are especially relevant to our understanding of resource use, and the potential for resource conservation.

1.2 Key Policy Issues

Prices have emerged as a key policy instrument for managing environmental infrastructure. Pricing policy, of course, only exists where prices are not set by competition, and prices often cannot be set by competition because of the natural monopoly characteristics of urban and environmental infrastructure. Train (1991) defines natural monopoly as when "the costs of production are such that it is less expensive for market demand to be met with one firm than with more than one" (p.1). In the building of environmental infrastructure there are often network economies, or economies of scale, that allow a single firm or builder to provide additional resources at the lowest possible cost. Furthermore, the building of competing infrastructure networks in the dense urban environment often poses high costs of entry and exit for other providers, as well as externality effects. Regulation is therefore needed to align the goals of the single firm with those of society, as measured by maximum social welfare, which is why utilities often operate under public ownership, or under more intensive regulation than other industries. There are also undisputed public good benefits of environmental resources, such as clean drinking water for public health. As Train (1991, p. 3) summarizes neatly, "public utilities, which are usually natural monopolies, play an essential role in the nation's economy and constitute one of the most prevalent settings for regulations in the country. Electricity, natural gas, local phone service, waste disposal, cable television, and many other goods and services are provided by public utilities subject to regulation by local or state agencies".

Although the provision of environmental goods has long been subject to heavy regulation, it has been increasingly recognized by even non-economists that the correct pricing of environmental goods can result in higher levels of investment, extended services, and gains in efficiency from demand reduction and better allocation between users. In 1992 the United Nations declared that "water is an economic good" (Rogers et al., 2002).

Prices also have many unique advantages to recommend them as a policy instrument. They are cheap to deliver through existing billing systems, and they can be targeted to specific categories of water use and customers. Prices can also be used to communicate with consumers that are difficult to contact, coordinate or persuade. Compared to the commercial or industrial building sectors where a majority of local resources may be consumed by a few small firms that seek to maximize profits and minimize costs, in contrast, water consumption decisions are made by thousands or millions of individual consumers in the single-family residential sector who all may have different behavioral reasons for using or conserving water.

Prices may in fact be the most effective policy instrument to ensure cooperation and compliance with conservation goals. Renwick and Green (2000), and more recently, Olmstead and Stavins (2008), and Coleman (2009) all find evidence that pricing policies can be more effective and longer-lasting than other demand side management efforts such as public information campaigns, education, and voluntary or mandatory watering restrictions. Optimal pricing can also complement and encourage other conservation actions such as technology standards and appliance and fixture retrofitting (Timmins, 2003; Deoreo et al., 2001).

One particular pricing tool used by utilities are nonlinear price structures, also commonly known as block rate structures, tiered pricing schemes, or multi-part tariffs. Generally, they get their name from the fact that marginal prices are not linear with respect to the quantity of consumption; instead, different prices are charged according to different blocks or tiers of consumption. These can be *declining* or *inverted* block rate price structures, in which prices decrease with higher incremental levels of consumption, or *inclining* or *increasing* block rate structures, in which increasing prices are charged. If you think about it, such nonlinear prices occur often in everyday situations, such as in the increasing prices of cellphone plans when you go over a particular limit; in decreasing prices in bundled goods to encourage consumption, such as in ‘buy two get one free’; or in the case of earnings, in the increasing (decreasing) marginal tax rate in progressive (regressive) taxation systems.

Utilities often used declining block rates in the early period of their development to encourage higher levels of consumption of environmental goods and thereby encourage the development of infrastructure networks. More recently, however, utilities have turned to increasing block rate price structures to signal the high environmental cost of new supplies; Hewitt (2000) describes why particular water utilities choose particular pricing schemes. Increasing block rates can also be used to preserve access to, and equitable distribution of, resources that are considered to be critical to human life, by charging lower or negligible costs for some minimal level of access (as in subsidized low-income heating programs), and by charging higher costs to higher quantities that are considered to be discretionary.

The related academic literature on energy has recognized the importance of prices on demand since Houthakker (1951). In water demand, Howe and Linaweaver (1967) introduced marginal price as a variable for residential water consumption. These papers kicked off an extensive debate on how to properly model the nonlinear characteristics of the price structure, because of the inherent bias introduced into ordinary least squares (OLS) estimation when price and quantity are determined simultaneously. This is called the *simultaneity bias*, and this debate about how to account for it has been frequently inconclusive. As Deller et al. (1986) grumbles, "the literature appears to be plagued with inconsistencies and there is no evidence of a consensus. This is not surprising given the inherent differences in data used in empirical analyses. One researcher may not have the data required to develop the [techniques] used in a previous study" (p. 334). In addition, there has been significant disagreement over the results obtained from previous analyses because of possible *aggregation bias*, which is introduced by analyzing the aggregated behavior of all users within utility districts or cities, rather than modeling the response of individual households when faced with nonlinear price structures. The econometric issues involved in these biases will be reviewed in greater depth in Chapter 2. Perhaps because of this lack of consistent and appropriate micro-data, the water literature has never been able to consider the possibility that individual households may differ in their response to pricing policies.

There are many good reasons for individuals, households, and various groups to respond differently, that is to be different or *heterogeneous* in their response to particular policies. There are the obvious physical or engineering factors: households may have a different number and composition of individuals and appliances. However, Lutzenhiser et al. (1987) observed significant variations in energy use between households and sociodemographic groups, even when compared in similar housing stock. This research has opened up new fields of inquiry and possibilities. Social structures, such as who sees and pays the bill, may determine how members of a household choose to consume resources. Attitudes common to sociodemographic groups or neighborhoods may also play a role in how households choose to use resources. Furthermore, there may be individual psychological or perceptual barriers to the understanding of how resources are consumed, since the act of switching on a water- or energy-using appliance may be unconnected to the perception of consumption, which may be further unconnected to knowledge of the bill. Individuals also may differ in their willingness or understanding to take advantage of conservation technologies or behaviors.

However, obtaining the data to explore these issues is often difficult and expensive, since it typically requires extensive fieldwork and surveys. Individuals and groups may also be heterogeneous in ways that are otherwise observed. For example, previous research has suggested that the size of the house lot, and the extent of lawn and vegetation, may have a significant effect on outdoor water use during the summer. In addition, the size of the house, number of bathrooms, and household income may all affect the responsiveness of different users to pricing policies. Very few studies manage to obtain this full range of data.

Heterogeneity among users, though largely neglected until now, has important implications for the design, planning, modeling, and operation of infrastructure. Although pricing policies are usually applied uniformly to maintain their simplicity, they can have distinct and complicated effects that vary between users and across geography. For example, although users may consume different amounts of water consistently, and pricing policies may affect them differently, the majority of infrastructure is built according to engineering codes that rigidly specify minimum flow and pressure requirements at every node. Increasingly sophisticated techniques have been used to design water distribution networks, but always assuming a uniform set of water delivery requirements (see, for example, Savic and Walters, 1997; da Conceicao Cunha and Sousa, 1999; Eusuff and Lansey, 2003). Furthermore, water systems in the United States are rapidly aging and failing, and will require massive initial and fixed cost investments to replace them (Duhigg, 2010). Modeling and predicting the use of water accurately in cities is critical to planning and designing efficient replacement networks. Finally, it is important to analyze the distributional implications of pricing policies on heterogeneous social and economic groups.

1.3 Research Contribution

The overall contribution of this dissertation is therefore to address the issues of *simultaneity*, *micro-data*, and *heterogeneity* in an integrated fashion. This dissertation uses a rich, new source of observational microdata obtained from SPU sampled from their billing databases. The block rate price structure combined with the significant policy change in 2001 constitutes an appropriate quasi-experimental design or 'natural experiment' that only affected high water users, since lower tier users were never affected by the higher price change.

In order to address the issue of simultaneity, I extend the various instrumental variables (IV) techniques suggested by the literature. By applying these IV techniques to the appropriate-level microdata in Chapter 4 using the correlated random coefficients (CRC) model, this dissertation obtains population-level estimates for price and income elasticities for heterogeneous households and neighborhoods.

In Chapter 5, using a hierarchical linear model and Bayesian estimating techniques, I extend the classical techniques used in Chapter 4 to explain the origins of these variations among households and groups in terms of observable and unobserved heterogeneity, and obtain estimates for the relative influence of geographic, physical, and socioeconomic factors on responsiveness to price. I also use recent developments in causal inference to obtain estimates of the local average treatment effect (LATE).

Finally, in Chapter 6, I estimate the impacts to welfare for the pricing policy change. The calculation of the Marshallian consumer surplus is contrasted and compared with other statistical techniques commonly used in applied economic analysis.

1.4 Context & Broader Goals

This dissertation was planned from the outset to build on my interest in environmental planning and water issues, as well as key research trends. I will first discuss the increasing importance of water issues, and then discuss how this research builds on several other emerging research trends.

This dissertation seeks to add quantitative techniques and analysis to address growing, multidimensional concerns about water issues worldwide. This alarm has not only reached the academic literature, but also the general press in publications such as *The Economist, The New Yorker,* and the *New York Times Magazine* (Peet, 2003; Specter, 2006; Gertner, 2007; Grimond, 2010). Continued urbanization, population growth, and development are all expected to strain existing water supplies further. In the developing world, the United Nations declared 2003 the "international year of freshwater", and the 2006 Human Development Report explicitly identified lack of clean water as perhaps the most important issue facing developing countries today (Watkins et al., 2006).

Water has always been critical for urban development, and in an era of increasing scarcity, it is therefore likely to emerge as a key integrative issue in urban planning practice and theory. For example, cities throughout the United States face issues of water supply: not only fast-growing cities in the arid southwest such as Phoenix and Las Vegas, but also cities on the eastern seaboard such as New York City, Washington, D.C., and of course, Atlanta (Stauffer, 2004; Johns, 2010). Perhaps the greatest stresses are emerging in rapidly developing countries such as China, which already faces severe water pollution and supply problems, combined with explosive urbanization, with approximately 750M people relocating to cities over the next thirty years (Yardley, 2004; Economy, 2004).

These problems will only be exacerbated in the future by climate change. Although water is a potentially infinitely renewable resource if managed properly, less than 2% of the world's water is drinkable (Smil, 2003). Of this, much of the world's water supplies are already claimed for use, and significant risks have been predicted in climate change scenarios at the local, regional, and global level (see, for example, Knowles and Cayan, 2002; Barnett et al., 2005; Postel et al., 1996; Postel, 2000). There is also a growing awareness of the linkages between energy and water, links that are likely to be strained by the growing water needs of electricity generation from alternative energy sources (Webber, 2008; Hoyle, 2008).

In order to contribute to the understanding of these concerns most effectively, this dissertation is intended to build strategically on several emerging research agendas in both the academic and policy literature. First, municipal governments have begun to play an increasingly large and vital role in the coordination and execution of environmental policies. Second, growing amounts of microdata allows the description of the heterogeneity that is characteristic of urban populations. Third, microeconometric models have emerged that allow the interpretation and synthesis of this microdata. Fourth, increasing attention to the issue of causal inference has enabled researchers to develop a better understanding of what can and cannot be learned from observational data, and how to look for appropriate experimental settings. Fifth, and finally, models that incorporate heterogeneity and microdata have begun to play an increasingly important role in the evaluation of policy for both equity and efficiency, both *ex ante* and *ex post*.

This dissertation is intended to take advantage of the research opportunities corresponding to these trends. The increasing role of municipal governments in environmental policy has created a significant need to model consumption at the urban level, an area that has not been previously explored because of data limitations. Growing amounts of microdata are now enabling researchers and policymakers to describe better heterogeneous populations and the distribution of preferences within them. Microeconometric models make it possible to test existing theories of behavior and to make quantitative statements about the relative importance of causes within theories. The rational evaluation and design of policy requires an assessment of the various distributional effects of policies and judgment whether the policy has reached its intended goals in an efficient manner.

The first major trend is the emergence of environmental policymaking at the local level. Over the past ten years, many cities in the United States have begun to pursue local environmental initiatives. The West Coast cities of Santa Monica, Portland, Seattle, and San Francisco are generally credited with some of the first municipal environmental efforts. This trend has further accelerated, with major cities such as Chicago, Boston, Los Angeles, and New York all pursuing major environmental initiatives, and many smaller cities pursuing similar policies (Portney, 2003). In recent years, international environmental efforts such as the Clinton Climate Initiative and C40 have been initiated and led by coalitions of large international cities.

There are several reasons why cities have emerged worldwide as particularly important foci for environmental action and concerns. First, the processes of urbanization and natural resource consumption are closely linked. As is often noted, the majority of humanity now lives in cities, and it is clear that as societies become more urbanized, they also consume more resources such as land, material, food, water, and energy (Berry, 1990; Vitousek et al., 1997). Second, in the absence of national and international policies on key issues such as climate change, mayors have emerged as key environmental advocates, and urban governments have therefore begun to pursue local environmental initiatives more aggressively (Kousky and Schneider, 2003). Third, although cities often constitute the smallest unit of political action, they still play a particularly significant role in determining how natural resources are used within their jurisdictions. As the United Nations declared at the Earth Summit in Rio de Janeiro, "local authorities construct, operate and maintain economic, social and environmental infrastructure, oversee planning processes, establish local environmental policies and regulations, and assist in implementing national and subnational environmental policies." (Nations, 1992, section 28.1).

Increasing resource flows into cities are often taken as a key indicator of the impact of cities on the environment, and concern with aggregate resource consumption is a common theme among municipal environmental policies. Focusing on the flows and changes in specific natural resources such as energy and water enables us to make conceptual connections between many disparate scales and systems, such as between individual consumption and aggregate environmental impacts; between technology, infrastructure, and behavior; as well as between social and ecological systems (Rees, 1992).

Despite this widespread interest in urban environmental policies, however, rigorous analysis of the effectiveness and impacts of local environmental policies is often difficult for a number of reasons. The meaning and measurement of 'sustainability' itself remains fiercely contested (Dryzek, 2005). In addition, the institutions necessary to study local environmental policy are still in the process of formation, often in municipal sustainability initiatives. At the moment, many environmental initiatives are initially justified either on a normative basis ("it's the right thing to do") or a technological-engineering basis ("new technology will solve this problem"). Local environmental policies, particularly within cities, still need to be developed to address diverse and complex issues, multiple scales in space and time for analysis, and dynamic changes. Most importantly for the relevance of this dissertation to policy analysis and evaluation, it has previously been difficult to develop appropriate models and theory that realistically describe the heterogeneity and complexity of the city, and to obtain data to test these theories.

In response to this need, the second major research trend are the opportunities made possible by the accelerating power of computers and the growing pervasiveness of digital information. The increased use of computerized systems for sensing, transactions, and storage – such as product scanners in grocery check-out counters, utility meter networks that monitor home energy usage, and intelligent traffic systems that monitor and shape traffic patterns – is rapidly creating large new databases of observed individual behavior.

The growth in available microdata is strongly reminiscent of two previous "statistical" revolutions in the social sciences. First, after World War II, many new kinds of national-level data became available, stimulating work in macroeconomics, demography, sociology, and other fields. Later, in the 1960's, growing amounts of survey information on individual behavior became available to researchers. Combined with the initial spread of computers, it became possible to analyze many new problems in many fields, including public health, market research, transportation, and economics, to name just a few.

New technology is also making existing techniques and methods more accessible and effective. Surveys are already a fundamental feature of research in many fields, but new information is continually being collected at an accelerating rate. Improvements in computer storage and communication makes it possible to build and disseminate digital information rapidly. Computer processing power has greatly increased, and many new analytical tools have been developed that take advantage of this, particularly in computational statistics. Furthermore, database tools and geographic information systems allow previously separate information to be synthesized and analyzed. Social networking and location-based services did not exist five years ago, but seem likely poised to generate even more data about behavior in particular settings.

The third major theme is the use of microeconometric models as a tool to study the growing sources of microdata. The importance of microeconometrics was validated by the 2000 Nobel Prize in Economics shared by James Heckman and Daniel McFadden. In their prize lectures, both economists describe microeconometrics as the fruitful combination of new available data, theories, and computational capabilities. Above all, both describe the goal of microeconometrics to describe more realistically the choices faced by heterogeneous individuals with different preferences. As McFadden writes,

> "Before the 1960's, economists used consumer theory mostly as a logical tool, to explore conceptually the properties of alternative market organizations and economic policies. When the data was applied empirically, it was to market-level or national-accounts-level data. In these applications, the theory was usually developed in terms of a *representative agent*, with market-level behavior given by the representative agent's behavior writ large. When observations deviated from those implied by the representative agent theory, these differences were swept into an additive disturbance and attributed to data measurement errors, rather than to unobserved factors within or across individual agents." (Mcfadden, 2001, p. 330).

The incorrect aggregation of data is a problem that has long been recognized to result in theoretically incorrect and empirically implausible estimates in many fields, for example in economics (Theil, 1954; Green, 1964; Fisher, 1969), in transportation (Koppelman, 1974), and specifically in water demand modeling (Danielson, 1979). Recent advances in microeconometrics takes advantage of more appropriate data to estimate more realistic models. As Heckman writes,

> "Microeconometrics extended the Cowles theory [of econometrics] by building richer economic models where heterogeneity of agents plays a fundamental role and where the equations being estimated are more closely linked to individual data and individual choice models. At its heart, economic theory is about individuals and their interactions in markets or other social settings. The data needed to test the micro theory are microdata." (Heckman, 2001, p. 256)

Microeconometric theory and microdata continue to find new applications throughout the social sciences.

The fourth major theme is that researchers have paid increasing attention to questions of causal inference from observational data. There is now a general consensus about a unified theory of causation, which has led to a well-articulated framework for understanding the opportunities and limitations of observational data (see, for example, Pearl, 2000; Morgan and Winship, 2007; Angrist and Pischke, 2009). This attention is particularly important given the explosion in observational data obtained in non-experimental settings.

The fifth and related major theme that this prospectus will build on is the development of microeconometric theories and microdata specifically to evaluate the effectiveness of existing policies (*ex post*) and to evaluate prospectively the implementation of policies in new environments (*ex ante*). Heckman (2001, p. 259) argues that a major development in microeconometrics has been the careful separation of these two themes, between the study of treatment effects, such as "what is the effect of a program in place on participants and nonparticipants compared to no program at all or some alternative program?"; versus questions of program design and implementation, such as "what is the likely effect of a new program or an old program applied to a new environment?". Questions such as the first are already staples of many disciplines of social science, formalized in texts such as Rossi et al. (2004). Structural models that take their structure from economic theory, yet are flexible enough to be applied to microdata, have increasingly begun to link these two questions into a broad framework to evaluate policy in a wide range of contexts, such as evaluating of existing policies; projecting the effectiveness of existing policies in new policy environments; forecasting the effects of new polices that did not previously exist; comparing parameters across empirical studies; testing of different economic theories; and making quantitative statements about the relative importance of causes within a theory. All of these features are not only desirable from the perspective of academic research and understanding, but are also of fundamental importance in policy design and implementation.

Chapter 2

LITERATURE REVIEW

2.1 Major Issues in Previous Work

A large body of academic literature already exists that describes the modeling of water demand. Review articles such as Hanemann (1998), and more recently, Arbues et al. (2003), comprehensively summarize more than 50 and more than 150 studies on water demand, respectively. These reviews categorize the previous studies according to the type of data used, the unit of analysis, variable and model specifications, and the general findings. The first review by Hanemann emphasizes how water demand modeling fits in the context of economic theory, with an emphasis on functional forms, while the second review by Arbues lists more comprehensively various techniques and specifications used in previous work.

One of the key debates over econometric technique has been over the specification of price. Price and income elasticities have been a key focus of the water demand literature, because of the policy importance of pricing policies, as discussed in the introduction (Chapter 1, Section 1.2). However, because of the particular estimation issues introduced by the nonlinear, or block-rate, price structures when observed marginal price is used, it is often possible to obtain positive price elasticities, which do not conform with economic theory or our intuition. I will now detail the first attempts to fix this problem.

After the initial inclusion of marginal prices in the modeling of demand for energy (Houthakker, 1951) and water (Howe and Linaweaver, 1967), the first attempt to deal with the nature of nonlinear price structures was the comprehensive survey of electricity demand by Taylor (1975). This paper raised many of the main issues in the modeling of electricity demand, and suggested that proper modeling of energy demand include both marginal and average prices. The subsequent modification by Nordin (1976), however, has found far greater application. Nordin proposed the inclusion of an additional "difference" variable to represent the effect of block rates as a lump-sum effect on income. In the case of decreasing

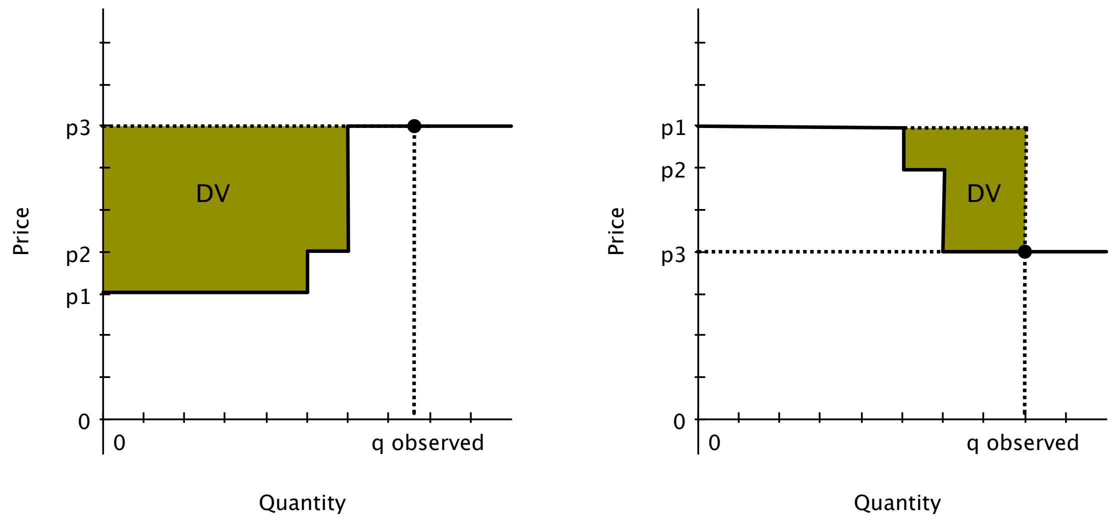

Figure 2.1: Calculation of the Nordin Difference Variable. The difference variable, or intra-marginal income transfer, is the shaded area for a consumer observed in the third block of an increasing block rate price structure (left), and a decreasing block rate price structure (right), respectively.

block rates, a customer observed in a higher block would need to pay an additional amount not represented by the observed marginal price. In the case of increasing block rates, the Nordin difference variable is computed as the gap between what the consumer would pay at the marginal price at quantity Q, compared to the amount they would actually pay under the graduated block rate structure. A rough sketch of the difference variable as an adjustment to income is shown below in Figure 2.1.

The water demand literature has largely debated the implications of the Taylor and Nordin specifications. Billings and Agthe (1980) provide a theoretical justification for using the marginal price and difference variable specification together. Despite widespread knowledge of the Nordin's proposed modification, numerous papers in the water demand literature still sought to use *either* marginal or average price because the individual household level data was not available to implement the difference variable specification. Specifying price as the marginal prices at the observed level of consumption leads to possible simultaneity biases, because price and quantity are simultaneously determined in the nonlinear price structures (Howe and Linaweaver, 1967; Jones and Morris, 1984; Pint, 1999). Specifying price as the average price introduces problems with causality, since the average price can only be calculated *ex post* by the consumer (Foster and Beattie, 1979; Griffin et al., 1981).

Other papers sought to conclude this debate by empirically testing whether consumers respond to marginal *or* average prices. Opaluch (1982, 1984) first devised a test to determine whether marginal or average prices are correct, followed by Chicoine and Ramamurthy (1986) and Nieswiadomy and Molina (1991). The average price specification has been subsequently argued for based on the lack of price knowledge among consumers, although the empirical results in the water demand literature remain inconclusive (Foster and Beattie, 1981; Shin, 1985).

The difference variable, however, in Taylor's theory should be equal in magnitude and opposite in sign to the price elasticity. Unfortunately, this also has been tested in the literature with little success (Billings and Agthe, 1980; Billings, 1982; Agthe et al., 1986; Nieswiadomy and Molina, 1989; Hewitt and Hanemann, 1995; Renwick and Green, 2000). Interestingly, Dalhuisen et al. (2003) finds this to be true in a meta-analysis (discussed further below).

Instrumental variables are the natural choice to solve problems with simultaneity or endogeneity. Many of the papers cited above also offer various instruments to compensate for the simultaneity inherent in the block-rate price structure (Jones and Morris, 1984; Agthe et al., 1986; Chicoine and Ramamurthy, 1986; Nieswiadomy and Molina, 1991). Many instruments have also been proposed to estimate the effect of multi-block price structures, driven by the absence of exact price or consumption information, as summarized in Deller et al. (1986); they assess the literature of the time in an even-handed manner when they write, "one technique of instrumental variable development can not always be preferred over another. The best approach depends, in part, on the nature of the available data" (p. 345).

It emerges as a common theme throughout the literature that data limitations have prevented numerous debates from reaching consensus. In addition to Deller et al's observation that techniques often remain un-compared, two other authors make convincing arguments that the aggregated nature of the data often used in the water demand literature is not appropriate for the conclusions frequently drawn. Danielson (1979) notes that the aggregated data frequently used in cross-sectional studies ignore changes in variables over time, and that the appropriate data would be for household-level time series data. Schefter and David (1985) observe that the aggregated nature of the datasets means that only the mean marginal price elasticity and mean difference variable can be obtained from aggregate data, and that these depend strongly on the variance within aggregated groups. They also suggest that household-level data would eliminate this source of aggregation bias. This echoes observations in the more general economic literature and in transportation (Theil, 1954; Green, 1964; Fisher, 1969; Koppelman, 1974).

Given the wide range of techniques, data, and contexts for many of these studies, perhaps the best summary estimates for price and income elasticities come from meta-analyses of the water demand literature. Espey et al. (1997) analyzes 124 estimates obtained from 24 articles, and excluding positive price elasticities, they find a median short-range price elasticity of -0.38, within a range from a minimum of -0.03 to a maximum of -2.23; and a median long-term price elasticity of -0.64 within a range from a minimum of -0.10 to a maximum of -3.33. Dalhuisen et al. (2003) follows up with a larger sample of 296 estimates from 64 studies, finding a sample median of -0.35, a sample mean price elasticity of -0.41, with a standard deviation of 0.86. Most of the price elasticities that they are find are negative, which conforms with theoretical and intuitive expectations. Dalhuisen et al. also find income elasticities with a sample mean of +0.43, a sample median of +0.24, and a standard deviation of 0.79.

One technique that has been applied to relatively few papers is the discrete-continuous choice (DCC) model. Arbues et al. (2003) note that this model has been infrequently used due to its expensive data requirements. However, Hewitt and Hanemann (1995) and Pint (1999) apply this model to household-level data, estimated by maximum likelihood, to obtain interesting if contradictory results. The former obtain very high price elasticities (-1.59) compared to the literature using OLS and IV methods, while the latter finds very low elasticities for the average user (approximately -0.09).

2.2 Heterogeneity

Very little work has been done that accounts for heterogeneity among water consumers. Only two papers in the broad water literature explicitly consider the wide observed variation among water users. Murdock et al. (1991) considers the role of sociodemographic characteristics in projections of water use, and Schneider and Whitlatch (1991) considers development of specific price elasticities for specific *user classes*, such as by commercial, industrial, residential users, and so on. The empirical conclusions of both papers, however, are frustrated by the same aggregate-data limitations cited above that plague much of the broader water demand literature.

More recent work, however, has sought to incorporate heterogeneity into the estimation of water demand, in particular using fixed and random effects in the intercept to set different levels of consumption for individual households. These more recent papers have been greatly facilitated by the use of household-level panel data. Pint (1999) compares a fixed effects model to her DCC model to account for household effects. Although the inclusion of household-level effects greatly increases the fit of her model, this specification does not include any IV specification, giving a positive price elasticity which leads her to discount this validity of this model. Arbus et al. (2004) allows individual variation within a dynamic panel data approach. Hanemann and Nauges (2005) use interaction terms to determine heterogenous responses to various price and non-price conservation programs in Los Angeles. Nataraj and Hanemann (2008) include fixed effects in a regression discontinuity model to account for household variation. Similarly, the recent papers by Coleman (2009) and Polebitski and Palmer (2009) both estimate fixed and random effects for individual households and census tracts, respectively.

All of these previous papers typically allow heterogeneity to occur in the intercept, but not in the slopes or coefficients. I can find no previous work considering the variation of price elasticities among heterogeneous consumers.

The most relevant techniques comes from labor economics, which consider the selection bias of education when estimating the causal effect of education on earnings. This is called the correlated random coefficients (CRC) model. Notable empirical papers in this field include Garen (1984) and Card (2001). Notable papers on the econometric issues involved include Wooldridge (1997); Heckman and Vytlacil (1998); Wooldridge (2003). I will apply this model to water demand in Chapter 4.

Multilevel modeling also allows for heterogeneity among individuals and groups, by allowing the regression modeling of data that is clustered in groups, and for which coefficients can vary by group. It is natural to apply this to urban problems, because specific effects are often clustered within a geographic area, such as within a neighborhood. Multilevel models have a long history in agriculture and education, "where it is natural to model animals in groups and students in classrooms" (Gelman and Hill, 2007, p. 276). This text, as well as Gelman et al. (2004), offer an extensive introductions to Bayesian approaches to multilevel models. This approach is relevant because it allows group-level data to be associated with individual observations as a predictor of group-level heterogeneity. This approach will be demonstrated in Chapter 5.

The results in the subsequent chapters will show that although mean *population-level* elasticities report in the literature are consistent, there are significant underlying differences between groups and neighborhoods. This has important implications for the assessment of pricing policy impacts, as well as distribution and operation of infrastructure networks.

2.3 Statistical Properties of Welfare Measures

There is already an exhaustive literature on the use of welfare analysis in policy analysis. Welfare analysis has been one of the main concerns of economics: interest in how to define, analyze, and influence social welfare runs throughout the history of economics in the work of Smith, Mill, Walras, Marshall, Pigou, Pareto, Schumpeter, Kaldor, Hicks, and many others (Blaug, 1996; Slesnick, 2008). Research in welfare analysis touches almost all areas of public policy analysis, including but certainly not limited to, benefit-cost analysis and valuation of environmental resources (Weimer, 2008; Freeman, 2003). In a review article, Slesnick (1998) writes, "the measurement of welfare forms the foundation of public policy analysis. A full consideration of taxes, subsidies, transfer programs, health care, reform, regulation, environmental policy, the social security system, and educational reform must ultimately address the question of how these policies affect individual well-being."

Yet considerable adaptation is still needed in applied work in order to connect welfare analysis to empirical work. Slesnick continues on, "while centrally important to many problems of economic analysis, confusion persists concerning the relationship between commonly used welfare indicators and well-established theoretical formulations".

In environmental economics, the literature generally seems to agree that Bockstael et al. (1989) were the first to point out that welfare measures are nonlinear functions of random variables, with challenging statistical properties that were previously unaddressed by the literature. A number of papers then followed seeking to quantify and correct the various biases in the welfare measures and their uncertainty. Adamowicz et al. (1989) analyze the variance of welfare measures calculated from parameter estimates of recreational travel cost data. Using a Monte Carlo analysis, they found that the variance of consumer surplus measures varied widely depending on the functional form chosen for the initial estimation.

Kling and Sexton (1990) also found that implausible welfare estimates often resulted when point estimates from demand functions were used. In particular, the linear and semilog form of demand – which are the most popularly used – resulted in welfare measures with either ratios of estimated parameters, or estimated parameters in the denominator. As they note, both of these situations can result in poor or uninformative welfare measures, first because the mean and variance of ratios of random variables are not well defined, and second, because estimates of parameters in the denominator close to zero can result in welfare measures or their associated errors blowing up. The solution that Kling and Sexton (1990) offered was to apply bootstrapping while using inequality restrictions to eliminate 'aberrant' estimates of the welfare function; in effect, they impose the welfare measure to be limited to values. An alternative suggestion that they propose, but do not explore further, are inequality restrictions or truncation applied through Bayesian priors. This will be further discussed in Chapter 6.

Various authors have also proposed ways to approximate the statistical properties of welfare measures. Bockstael and Strand (1987) examine the interaction of multiple sources of error in welfare measures. Krinsky and Robb (1986) offer a simulation method to approximate the statistical properties of elasticities, by drawing from distribution with estimated parameters in order to *generate* empirical distributions of the elasticities, and Layton and Brown (2000) use this to obtain confidence intervals for welfare measures (in this case, WTP). Kling (1991) compares the relative precision of welfare measure error estimates from a Taylor Series approximation, the bootstrap, and the Krinsky-Robb method. Analytical alternatives are explored further in Chapter 6.

Chapter 3

DATA

The dataset used for this paper is a new, rich panel dataset of water consumption, obtained from the billing databases of SPU. This first two sections of this chapter will describe how additional data sources were added to the billing database, and how sampling was conducted for analysis. The last section of this chapter will describe the general features of the data to give additional insight into the results from Chapters 4 - 6.

3.1 Data Sources

The complete billing databases of SPU, the municipally-owned water and solid waste utility for the City of Seattle, were obtained for the period from July 1991 to December 2007. The data from before 2001 came from the "Waterbird" database, and after 2001 from the "CIDS" database. These billing databases were based on visual readings of water meters, occurring approximately on a bimonthly basis for most customers (the average time between meter reads is 53 days). These databases also include corrections to water meter reads. Most importantly, each meter reading is uniquely identified by a geocode and the utility account number. These two fields allow us to see the geographic location of the water consumption, and whether or not the customer at this location has changed from one period to the next. This is also a crucial advantage in joining this data to other data sources as well. The complete data cleaning process is described in the next section, Section 3.2.

Sources for the data joined to the water consumption information are summarized below in Table 3.1. Household level data originated with the 2007 King County Tax Assessor Database, which is maintained by the King County Department of Assessments, and this data was obtained from the Washington State Geospatial Data Archive, which is maintained by Map Collection & Cartographic Information Services, University of Washington Libraries (http://wagda.lib). Tract-level information was obtained from the 2000 U.S. Census. Climate data was obtained from the National Oceanographic & Atmospheric Administration (NOAA) data and downloaded from the National Climatic Data Center (Karl et al., 1986). The specific weather station used was the Seattle-Tacoma International Airport weather station, which although several miles from downtown, has the longest uninterrupted time series of weather data in the region.

Category	Source	Description	Abbreviation
Price	SPU	Average daily consumption	Q
		Marginal price	MP
		Difference variable	DV
		Price, block 1	P1
		Price, block 2	P2
		Price, block 3	P3
		Sewer, marginal cost	Sewer
		Monthly fixed price	Fixed cost
Climate	NOAA	Average daily temperature	Average temp
		Average precipitation	Average prcp
Household	King County	House value (2007)	Value
		Lot size	Lotsf
		House age	Age
Census tract	Census 2000	Median income	Med income
		Average household size	Av HH size
		Median number of rooms	Med rooms
		Median house value	Med value
		Population by tract	Pop

Table 3.1: Data Sources Joined to Original Billing Data.

3.2 Data Cleaning

This section will narrate the process by which the data was cleaned and prepared for analysis. Extensive data cleaning was necessary, since the billing data was not originally collected for observational purposes, and other data sources had to be joined to it.

The SPU billing databases contain 18.45 million records in the period between 1991 and 2007. Each record, in addition to its unique identifying information, contained the quantity of consumption in ccf, the meter read date, the number of days between each meter reading, and the billing class code. Only single-family residential units were selected, leaving out duplexes, townhomes, and multifamily apartment buildings. Since the water meters are only read periodically, it is only possible to know the average daily water consumption over the billing period by dividing the total water consumption by the number of days between readings. The average daily consumption over a month was also calculated by dividing the total level of consumption by the total number of days in the billing period and multiplying times 30.5 days.

First, consistent geocoding was placed across all years and accounts. The billing accounts indicated by the geocodes were matched to King County tax assessor parcels using the account numbers, SPU's unique premise identification number, and the 10-digit King County parcel identification numbers. Some of the billing accounts did not have proper geocoding, and approximately 1% of the data was eliminated this way. In addition, a Geographic Information System was used to place each parcel within a census tract, and to obtain a census tract identification number. There were 124 residential census tracts within the City of Seattle in the 2000 Census. Household-level data was then joined from the 2007 King County Tax Assessor database. Since historical data could not be obtained for the full period of time for which there were bills, I used only the 2007 database.

Time-varying climate data was then joined to the individual billing records. The average daytime temperature for each of the days, over the period of time represented by the billing record, was joined to each individual billing record. The total precipitation for each of the days, over the period of time was represented by the billing record, was also joined to each individual billing record. This also made it possible to calculate the average daily precipitation over the time of each billing record.

Finally, the rate history information obtained from SPU was joined to the billing records (Flory, 2008). Figure 3.1 below shows the price histories for each block, as well as the sewerage and monthly fixed charges. Because each meter reading only contains the total water consumption and number of days since the past reading, SPU calculates the block rate charges based on average water use over a month. It is therefore possible to calculate for each meter reading all of the following: the total bill; the highest tier of consumption and marginal price; the quantities and prices for each lower block. Therefore the monthly fixed cost of connection to the system, the price in each block, and the current sewer rates

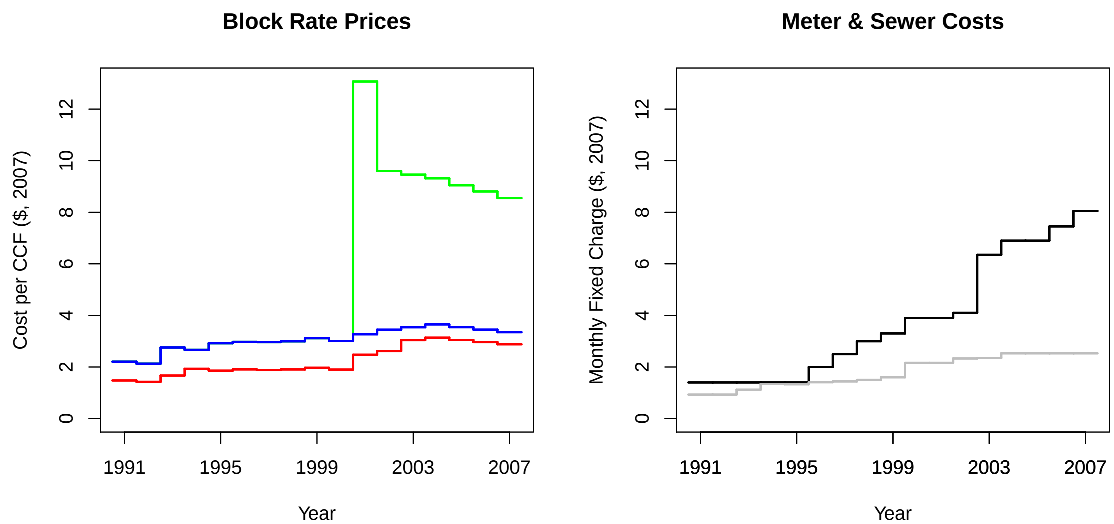

Figure 3.1: SPU Water Price History. On the left, the red and blue lines represent the first and second block prices for consumption between 0-5 ccf, and 5-18 ccf, respectively. The green represents the third price tier only imposed on consumption greater than 18 ccf after 2001. On the right hand graph, the gray line represents the monthly fixed price and the black line is the sewer rate.

were joined to each billing record. Taking the total quantity of consumption for each bill, and knowing the thresholds for each price block in each year, allows the total bill to be calculated, as well as the marginal price observed at level of consumption, the average price paid, and the total bill. The difference variable was also calculated as the total bill minus the bill if all of the consumption had occurred at the observed marginal price (see Figure 2.1). For an inclining block rate system such as Seattle's, all difference variables are zero or negative, because if all consumption would be measured at the marginal price observed, then the total bills would be higher than were actually paid.

Sewerage rates are sufficiently complicated to require additional discussion. SPU charges sewerage rates based on the last year's consumption, basing their assessment of last year's consumption on the lower of the winter water consumption or the summer water consumption. For the vast majority of users and circumstances, this is based on the summer consumption, which is typically much higher. The annual total sewer bill is then spread throughout the year. This rule was sufficiently complicated that it could not easily be modeled, and I simply took the current year's sewer rates in the analysis. This seems to be a reasonable approximation, considering the difficulty that an actual customer would have to remember last year's sewer rate and consumption levels.

All water and house prices were adjusted by the U.S. Consumer Price Index in 2007 real dollars. Additional assumptions were made in the sampling process, which is described in the next section.

3.3 Sampling Method

A balanced sample was created by extracting approximately 100 accounts from every census tract. Although there are more than 120 census tracts in Seattle, only 97 census tracts have more than 1000 household accounts, and so the analysis was limited to these 97 tracts. These 97 tracts constitute 93% of all water accounts in the database and represent the most densely populated areas of the city.

For the purposes of this study, only meter readings during the peak summer season were used, because this is when SPU applies an increasing block rate price structure. The peak season was extended from 3 months to 4 in 1995. Meter reading periods, when they span the change between peak and non-peak periods, are broken into two different meter readings prorated by the number of days in each type of period.

In addition, outlier records were omitted. Water bills with less than 15 days or greater than 150 days between readings were omitted from the data, because they likely represent anomalous readings by the field crew either for the purposes of bill investigations or corrections. Accounts with only more than 50 readings were used, in order to provide enough data to obtain account-level estimates for intercepts and coefficients in Chapters 5 and 6.

These sampling rules yielded 8,741 unique household accounts spread over 97 tracts. Sampling without replacement yielded 190,236 total observations, or approximately 1% of the original billing database. Three smaller sub-samples were then drawn from the original sample to use in smaller tests or cross-validation, by drawing exactly five accounts for each census tract. Descriptive statistics for the datasets appear below in Table 3.2.

Summary Variable	Overall Sample	Subsample 1	Subsample 2	Subsample 3
Observations (number)	190,236	10,605	10,366	10,609
Accounts (number)	8741	485	485	485
Tracts (number)	97	97	97	97
Average consumption per bill	13.687	13.587	13.36	13.566
Average read days per bill	48.685	48.789	48.78	48.756
Average daily consumption (ccf)	0.269	0.267	0.261	0.265
Median total bill ($)	17.65	17.63	16.559	17.503
Median marginal price paid ($)	3.043	3.043	3.043	3.043
Median difference variable ($)	-2.49	-2.49	-2.49	-2.49
Median price in block 1 ($)	2.616	2.616	2.616	2.616
Median price in block 2 ($)	3.447	3.447	3.447	3.447
Median price in block 3 ($)	8.807	8.807	8.807	8.807
Median sewer price ($)	5.828	5.828	5.828	5.828
Median fixed monthly cost ($)	4.634	4.604	4.604	4.604
Median house value ($, tract)	229,923	230,381	227,238	235,113
Median lot size (sf, tract)	5,500	5,684	5,350	5,300
Median house age (years, tract)	63.5	63	65	63
Median income ($, tract)	51,760	51,760	51,760	51,760
Median household size (people, tract)	2.201	2.191	2.197	2.201
Median number of rooms (in tract)	5.2	5.1	5.2	5.1
Median house value ($, tract)	259,900	256,400	256,400	262,000
Average temperature (degrees F)	63.642	63.629	63.632	63.642
Average precipitation (inches)	4.065	4.054	4.056	4.065

Table 3.2: Descriptive Statistics for Main Sample and Subsamples. Not surprisingly, most of the tract-level summary variables are the same, since a balanced sample was obtained across census tracts for approximately the same number of observations.

Chapter 4

CORRELATED RANDOM COEFFICIENTS MODEL OF WATER DEMAND

4.1 Introduction

As the previous chapters argued, very few studies in the broader water demand literature have acknowledged the possibility of heterogeneity among consumers. Although instrumental variable (IV) methods have been used to explore how consumers respond in complex pricing situations – such as estimating demand within the nonlinear, block-rate price structures typically used by utilities – no previous studies have explained how different consumers might respond *differently* to changes in pricing policies. This chapter introduces the correlated random coefficients (CRC) model in order to address the interaction between the simultaneity biases introduced by these pricing structures, and the inherent unobserved heterogeneity between consumers.

Unobserved heterogeneity has many possible sources in water demand. Many features that affect water demand cannot be observed without household survey information, which remains both difficult and expensive to obtain. Furthermore, it may be even more difficult to observe this heterogeneity consistently over time. Two otherwise similar households may have varying attitudes, appliance stocks, or household composition or numbers of occupants over time. In addition, there may also be effects that determine water consumption at the group-level, whether they are socio-demographically, economically, or geographically-defined. For example, a particular neighborhood may share certain physical characteristics such as the age and type of housing stock, or the level of outdoor vegetation. Social attitudes about water consumption may vary by individual or group, such as the amount of water that is acceptable to use for outdoor watering. Even the effectiveness of the water supply or sewerage functions may vary by neighborhood level, which in turn may affect how much water households may consume.

In the CRC model, the coefficients of endogenous variables, such as instruments or treatments, are considered to be random variables with the possibility of correlation. This allows the specification of models in which the intensity of the treatment varies among different individuals or groups. The random variable specification allows the instruments to interact with unobserved heterogeneity. Most importantly for the purposes of this policy evaluation, this model allows the coefficients of marginal price – i.e, the price elasticities – to vary across individuals and groups, such as by neighborhood or by census tract.

This chapter will develop the CRC model for water demand, and in doing so, seeks to answer the following key questions: First, what specification of the CRC model best fits the data at hand? Second, what do these different models tell us about the effect of prices on individuals and groups? Third, and finally, what is the average treatment effect of the change in marginal prices, and the pricing policies?

The next section will first introduce recent developments in the CRC model. Results are presented from the CRC model using the Terza IV method, which will be discussed further below. These are compared with results from the OLS, 2SLS, and random effect (RE) models that have previously been used in the water demand literature. This chapter will conclude with a comparison and contrast of the model results, and questions for further development and work.

4.2 Theory

This section will briefly describe the main issues in estimating the CRC model, to inform the subsequent model specification, estimation methods, and discussion.

The CRC model has been developed in response to problems in which the explanatory variables are often correlated with the observed outcomes. One area in which this model has been used repeatedly is in estimating the labor market returns to education (Garen, 1984; Card, 2001). Initial self-selection towards education in turn leads to higher incomes; the challenge is to disentangle the selection effects towards education and to measure the incremental income benefits of education. The CRC model also avoids multicollinearity introduced by ordinary least squares estimation with dummy variables (OLSDV), when using large numbers of indicators.

Theoretical issues in estimation of the CRC model are thoroughly explored in several papers. Wooldridge (1997) demonstrates that the 2SLS estimator consistently estimates the average treatment effect in a model with random coefficients, Heckman and Vytlacil (1998) consider the possibility of multiple treatments, and Wooldridge (2003) subsequently showed that standard IV estimators are consistent under relatively weaker assumptions. A short description following Wooldridge's treatment, adapted for longitudinal data, follows.

The outcome y_t in time period t is modeled in error form as:

$$y_t = \mathbf{a} + \mathbf{w_t}\mathbf{b} + e_t \tag{4.1}$$

where $\mathbf{a}$ and $\mathbf{b}$ are the vectors of random intercepts and coefficients, respectively, which may depend on covariates, unobserved heterogeneity, or their interactions.[1] $\mathbf{w_t}$ is the (endogenous) treatment in time interval t, and e_t is the error term. For a observation i in time period t, and with varying treatments over time and observations, denoted by $\mathbf{w_{it}}$,

$$y_{it} = \mathbf{a_i} + \mathbf{w_{it}}\mathbf{b_i} + e_{it} \tag{4.2}$$

Extending this notation for a observation i belonging to a group $j[i]$ – for example, multiple observations within a particular household or neighborhood – we can write the outcome as a function of the group-level intercept $a_{j[i]}$ and $b_{j[i]}$:

$$y_{it} = \mathbf{a_{j[i]}} + \mathbf{w_{it}}\mathbf{b_{j[i]}} + e_{it} \tag{4.3}$$

As Wooldridge writes, "because $\mathbf{b}$ is generally a function of unobserved heterogeneity, we cannot hope to estimate the vector of slopes, b for any particular cross-section unit, $j[i]$.

[1]Letters are used to emphasize the random variable nature of the parameter vectors $\mathbf{a}, \mathbf{b}$, though they could easily be generalized to a vector of fixed variables as well. This is the usual, confusing distinction between 'fixed' and 'random' effects. I will occasionally refer to them using the more specific names of 'varying intercept' and 'varying coefficient' models, though generally categorize them as 'random effect' and 'correlated random coefficient' models to compare results with the literature.

Instead we focus on the average treatment effect,

$$\beta \equiv E(\mathbf{b}) = E(\mathbf{b_{j[i]}}) \tag{4.4}$$

where the expectation is over the relevant population. The goal is to consistently estimate β given the availability of instrumental variables and sufficient exclusion restrictions." (Wooldridge, 2003, p. 187; original subscripts changed to maintain consistency). The goal of this chapter is to estimate the element of β that is associated with changes in the marginal price treatment, which is represented by an instrument.

Wooldridge (2003) shows that 2SLS with the endogenous instruments, the covariates, and interactions yields consistent estimates of β under the following key assumptions, similar to the usual exclusion restructions made for IV estimation: ignorability with respect to the covariates $\mathbf{x}$ and instrumental variables $\mathbf{z}$; redundancy of $\mathbf{z}$ for a and $\mathbf{b}$, conditional on $\mathbf{x}$, with linear conditional expectations; and that the conditional covariances do not depend on the instrumental variables.

4.3 Model Specification

Applying the general form of the CRC model to water demand, the standard 2SLS estimation for each individual or group j and each observation i can be written as follows:

$$\text{Stage 1:} \quad \hat{P_{ij}} \sim f(p_1 \ldots p_k, Z_j) \tag{4.5}$$

$$\hat{D_{ij}} \sim g(p_1 \ldots p_k, Z_j) \tag{4.6}$$

$$\text{Stage 2:} \quad Q_{ij} \sim h(\hat{P_{ij}}, \hat{D_{ij}}, Z_j) \tag{4.7}$$

where for the ith observation and jth household, $\hat{P_{ij}}$ and $\hat{D_{ij}}$ represent the predicted marginal price and difference variable instruments estimated in the first stage, respectively, as a function of $p_1 \ldots p_k$, which represent the block prices for blocks $k = 1 \ldots K$, and Z_j represents the other sociodemographic characteristics of the individual or group. Q_{ij} is the observed water consumption depending on the treatment intensity (as a function of predicted marginal prices and difference variable instruments, including covariates).

In the IV method used for the RE and CRC models, the Terza IV method is used. The Terza IV model obtains instruments by predicting the quantity as a function of the individual block prices, and then re-inserts the predicted quantities into the rate schedule to obtain an instrument for marginal price and the difference variable. This avoids the problem of more common IV methods which often predict marginal prices or difference variables that in fact do not exist within the nonlinear price structure. The stages of the estimation for the Terza IV would be instead written as:

$$\text{Stage 1:} \quad Q_{ij} \quad \sim f(p_1 \ldots p_k, Z_j) \tag{4.8}$$

$$\text{Stage 2:} \quad R(\hat{Q_{ij}}) \quad \rightarrow \hat{P_{ij}}, \hat{D_{ij}} \tag{4.9}$$

$$\text{Stage 3:} \quad Q_{ij} \quad \sim h(\hat{P_{ij}}, \hat{D_{ij}}, Z_j) \tag{4.10}$$

with the rate schedule written as a function R.

To expand this further into a CRC model, the intercepts and coefficients of the model can be represented as random draws from statistical distributions, which in turn may be parametrized differently for the appropriate groupings of data. This allows heterogeneity between groups by allowing selected model parameters to be drawn from differently parameterized statistical distributions. This also focuses attention on the parameters of the statistical distributions as a way to explain overall group-level effects.

In order to highlight the key variables in the subsequent development of models, we first write out the linear regression model below and then incrementally modify this linear equation in each of the subsequent models. The linear regression form for the average daily consumption within a month, Q, can be written as a statistical observation of a normal distribution, such that:

$$Q \quad \sim \quad \mathrm{N}(\mathbf{a} + \boldsymbol{\beta_i X_i} + b_P\hat{P} + b_D\hat{D}, \sigma^2) \tag{4.11}$$

where $\mathbf{a}$ is the intercept; $\boldsymbol{\beta_i}$ is the vector of fixed effects for observation-level predictors $\boldsymbol{X_i}$; b_P, b_D are the coefficients of the predicted marginal price and difference variables from Stage 1 ($\hat{P}$ and $\hat{D}$, respectively); and σ^2 is the residual variance. Now we must pay closer attention to the subscripts for our outcome variables and predictors. We will build up to the general CRC form by following the general model specification suggested by Laird and Ware (1982) for longitudinal data, in order to specify the correct groupings of data for each of the models along the way, including OLS, 2SLS, and RE models.

Pooled OLS model using observed price and difference variable: A completely pooled regression model that is regressed on all observations i can be estimated using ordinary least squares regression on the observed marginal price P and difference variable D. This would be written as:

$$Q \sim \mathrm{N}(\mathbf{a} + \boldsymbol{\beta_i X_i} + b_P P_i + b_D D_i, \sigma^2) \tag{4.12}$$

Pooled 2SLS model: A completely pooled regression model that regresses all observations i, using the two-stage instrumental variables technique, can be written as:

$$Q_i \sim \mathrm{N}(\mathbf{a} + \beta_i X_i + b_P \hat{P}_i + b_D \hat{D}_i + b_Z Z_i, \sigma^2) \tag{4.13}$$

where $\hat{P_{ij}}$ and $\hat{D_{ij}}$ represent the predicted marginal price and difference variable instruments estimated in the first stage as in equations 4.5 and 4.6, using a similar pooled specification in two separate regressions:

$$\hat{P}_i \sim \mathrm{N}(\mathbf{a} + \boldsymbol{\beta_i X_i} + \sum b_m p_m, \sigma^2) \tag{4.14}$$

$$\hat{D}_i \sim \mathrm{N}(\mathbf{a} + \boldsymbol{\beta_i X_i} + \sum b_m p_m, \sigma^2) \tag{4.15}$$

where $1 \ldots m$ are the individual tiers and p_m are the individual tier prices.

Random effect (RE) models: Allowing the intercept to vary by group j allows each individual to have a specific level of consumption:

$$Q_{ij} \sim \mathrm{N}(\mathbf{a}_{j[i]} + \boldsymbol{\beta_i X_i} + b_P \hat{P}_{ij} + b_D \hat{D}_{ij} + b_z Z_{ij}, \sigma^2) \tag{4.16}$$

$$\mathbf{a}_{j[i]} \sim \mathrm{N}(\mu_j, \sigma_j^2) \tag{4.17}$$

where $\mathbf{a}_j$ indicates a varying-intercept for each group j, drawn from a normal distribution which parameterized by a group mean μ_j and within-group variance σ_j^2.

CRC models: In the CRC model, we allow add a second level of random parameters, in which intercepts and coefficients are combined into a $1 \times K$ vector $\mathbf{b}_{j[i]}$ with a covariance matrix Σ^2 of dimension K for the number of covariates and instruments used. The observation level model is written as:

$$Q_{ijt} \sim \mathrm{N}(\boldsymbol{\beta_i X_i} + \mathbf{b}_{j[i]}\hat{\mathbf{w}}_{ijt}, \sigma^2) \tag{4.18}$$

$$\mathbf{b}_{j[i]} \sim \mathrm{MVN}(\boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j^2) \tag{4.19}$$

where $\hat{\mathbf{w}}_{ijt}$ is the model matrix that includes an initial column of 1's, allowing the first element of $\mathbf{b}_{j[i]}$ to describe the intercept previously denoted by $\mathbf{a_{j[i]}}$. Other columns include the variables for $\hat{P}_{ij}$ and $\hat{D}_{ij}$. In turn, the parameters of the multivariate normal, the group mean vector $\boldsymbol{\mu}_j$ and the covariance matrix Σ_j for each group j, allow correlations between the different variables.

4.4 Estimation Methods

All estimation was performed using the R language (Team, 2009, version 2.10.1). Linear regressions for the OLS model, was performed using the standard ordinary least squares function for linear regression, lm().

The 2SLS model was fit using the standard linear regression models which give consistent (unbiased) results, but the appropriate standard errors and t-values using the Wald estimator (Angrist and Pischke, 2009). Results were obtained using the tsls() function in the sem package (Fox, 1979, 2006).

The RE and CRC models were fit as linear mixed models using maximum likelihood using the lme4 package, developed by Bates (2005). A brief summary of the methods used by this package appears below, which is adapted from Bates (2010, chapter 5); see this reference for a more thorough formulation.

The linear mixed model is defined in terms of a q-dimensional vector of random effects, $\boldsymbol{B}$, and a n-dimensional response vector, $\boldsymbol{Y}$. The model is redefined in terms of the following distributions:

$$\begin{aligned} (\boldsymbol{Y}|\boldsymbol{U}=\boldsymbol{u}) &\sim \mathrm{N}(\boldsymbol{Z\Lambda_\theta u}+\boldsymbol{X\beta},\boldsymbol{\sigma^2 I_n}) \\ \boldsymbol{U} &\sim \mathrm{N}(0,\sigma^2 I_q) \end{aligned} \tag{4.20}$$

where $\boldsymbol{\Lambda_\theta}$ is a relative covariance matrix introduced to avoid a singular random vector covariance matrix (i.e., which is why the second line is not simply specified with a $q \times q$ covariance matrix $\boldsymbol{\Sigma}$). The linear predictor and conditional mean of the response are, respectively,

$$\gamma = \boldsymbol{Z\Lambda_\theta u}+\boldsymbol{X\beta} \tag{4.21}$$

$$\mu = E[\boldsymbol{Y}|\boldsymbol{U}=\boldsymbol{u}] \tag{4.22}$$

To solve for the parameters, we consider them to be varying and the observed data $\boldsymbol{y_{obs}}$ to be fixed. Solving for the marginal density of Y, and evaluating that density at $\boldsymbol{y_{obs}}$, gives us the parameter estimates that are most likely to result in our data. The joint density of $\boldsymbol{u}, \boldsymbol{y}$ at $\boldsymbol{y_{obs}}$ is:

$$f_{U|Y}(\boldsymbol{u}|\boldsymbol{y_{obs}}) = \frac{h(\boldsymbol{u})}{\int h(\boldsymbol{u})d\boldsymbol{u}} \tag{4.23}$$

where $h(\boldsymbol{u})$ is the *unnormalized conditional density*, and the integral in the denominator is the likelihood to be evaluated:

$$L(\boldsymbol{\theta},\boldsymbol{\beta},\boldsymbol{\sigma}|\boldsymbol{y_{obs}}) = \int h(\boldsymbol{u})d\boldsymbol{u} \tag{4.24}$$

We then use normal densities to form the continuous density $h(\boldsymbol{u}) \propto f_U \times f_{Y|U}$, where

$$f_{Y|U}(\boldsymbol{y}|\boldsymbol{u}) = \frac{1}{(2\pi\sigma^2)^{n/2}}\exp\left(-\frac{||\boldsymbol{y}-\boldsymbol{X\beta}-\boldsymbol{Z\Lambda_\theta u}||^{\boldsymbol{2}}}{2\sigma^2}\right) \tag{4.25}$$

$$f_U(\boldsymbol{u}) \quad = \quad \frac{1}{(2\pi\sigma^2)^{q/2}}\exp\left(-\frac{||\boldsymbol{u}||^2}{2\sigma^2}\right) \tag{4.26}$$

We then get

$$h(\boldsymbol{u}) = \frac{1}{(2\pi\sigma^2)^{(q+n)/2}}\exp\left(-\frac{\boldsymbol{||y - X\beta - Z\Lambda_\theta u||^2 + ||u||^2}}{2\sigma^2}\right) \tag{4.27}$$

Taking the deviance, or $-2 \times \log(h(\boldsymbol{u}))$, we get

$$-2 \times \log(h(\boldsymbol{u})) = (n+q)\log(2\pi\sigma^2) + \frac{\boldsymbol{||y - X\beta - Z\Lambda_\theta u||^2 + ||u||^2}}{2\sigma^2} \tag{4.28}$$

The objective function to be minimized with respect to u is therefore

$$\hat{u} = \arg\ \min_u \boldsymbol{||y - X\beta - Z\Lambda_\theta u||^2 + ||u||^2} \tag{4.29}$$

The first term is the *penalized residual sum of the squares* and the second term is a penalty on the length of the random variable vector, in order to provide a balance between the fit of the model to observed data versus the number of parameters needed.

Many alternative approaches were considered for an assessment of goodness of fit and model selection. ANOVA approaches do not work well to assess the various two-stage model specifications above, since the first stage models actually change the data in the second stage, making likelihood ratio tests impossible, since the second stage models are actually fit to different datasets. However, since all of the models are specified using the same fixed observation-level predictors $\boldsymbol{\beta_i X_i}$, the likelihood functions should be the same, and only the random effects added vary from model to model. Therefore, the Akaike Information Criteria, as a penalized measure of the deviance, was used to assess the likelihood of each model with one additional parameter for each random variable, calculated as

$$\text{AIC} = -2\log\text{likelihood} + 2 \times n$$

where n is the number of parameters used, so the smaller the AIC value, the better.

The most general approach for testing the accuracy of the predicted models is to use cross-validation of the models using two subsamples drawn from the data. The first sample can be used to obtain a model and then applied to the other model, and the accuracy of results can be assessed using root mean square error (RMSE).

4.5 Model Results

The results below are described in the sequence of models fit, with commentary on the various features of each model. The end of this section will discuss the relative model fits as measured by RMSE, and a comparison of the calculated price elasticities with the literature.

Pooled OLS model The OLS model was fit using the observed marginal price and difference variable as regressors. Estimated coefficients and standard errors for the estimated model are presented in the first column of Table 4.1. The data is center-normalized, so the magnitude of the coefficients can be compared in terms of standard deviations. The magnitude of the marginal price (+24.35) means that a one standard deviation change in price would have a very large effect on water consumption. Most importantly for the purposes of comparison, marginal price is positively signed (+24.35). This is inconsistent with theory and intuition, but is similar to the simultaneity bias and unexpected signs observed in other studies that use observed marginal price as a regressor.

Most of the other regressors have relatively small magnitudes, but show the expected sign and relationship with the observed daily average water consumption. Water consumption increases with increasing average temperature (+1.88) and decreases with increasing sewer costs (-3.54) and fixed costs (-7.26). The indicators for 1992 and 2001 both result in a decrease in water consumption (-0.37 and -2.59, respectively), likely reflecting the mandatory and voluntary water restrictions implemented in those years, respectively. The coefficient with the least intuitive interpretation is for average precipitation; this apparently increases water consumption with a positive effect (+0.11), though the standard errors is relatively large.

Overall, the fit of the model is mediocre. The adjusted R^2 value is 33.59%. However, Figure 4.1 plots the actual observed data versus the predictions, and reveals some of the flaws in the model. The model has clear horizontal lines which reflect the discrete step changes in over time. Since the higher-level observed values sit below the diagonal line, this model consistently under-predicts higher levels of water consumption. Finally, the model does a very poor job of predicting the very highest levels of consumption, which is of special interest because these users use much more than the average user and constitute a significant portion of water consumption.

Pooled 2SLS model The pooled 2SLS model was fit using a first stage model, in which the marginal price and difference variable are predicted from the individual block prices, and then puts these predicted quantities as instruments back into a second-stage regression. Results are shown in the second column of Table 4.1. The data is center-normalized, so the magnitude of the coefficients can again be compared in terms of standard deviations.

Some of the coefficients have changed when compared with the pooled OLS results in the same table. The fixed intercept remains large relative to the effect of the regressors. An increase in average temperature continues to show a strong positive effect on water consumption (+3.14). However, an increase in average precipitation now results in a decrease in water consumption as expected (-0.69). The coefficients for the fixed monthly cost (+0.30 with a standard error of 0.20) and the indicator variable for 2001 (+0.08 with a standard error of 0.11) are now relatively insignificant.

Most importantly, however, the use of the 2SLS method and instrumental variables leads to a negative coefficient for marginal price (-2.93), which is consistent with economic theory and our intuition.

Overall, the fit of the 2SLS model remains poor at the high end of water consumption. The residual standard error of the pooled 2SLS model (28.05) has increased markedly over the pooled OLS model (19.71). Figure 4.1 shows that the pooled 2SLS model achieves a high level fit by predicting most observations in a fairly tight range, where most of average water consumption occurs. However, the pooled 2SLS model still does not predict any of the highest water consumption observations accurately.

RE models Fixed parameter estimates from the RE models, grouped by account and tract, are presented in Table 4.2. For the intercept, indicated by an asterisk (*), the intercept value reported in this table is the mean, and the random effects are drawn from a normal distribution with mean zero. The variances are shown in Figure 4.2 and summarized in Table 4.4. The first half of Table 4.2 shows the fixed effects when a varying intercept is added for each account, and the second half shows the results when a varying intercept is added for each tract. Since the same data is used in both groupings, the only difference between the two fits is how many random variables are added to the model. When grouped by account, 8,741 random variables are added, one for each group. When grouped by tract, 97 random variables are added, one for each group. In each grouping, the RE model (also known as 'varying intercept') is fit first using the observed marginal price (models RE-1A and RE-1T), and then fit using the Terza instrument, as discussed above.

Moving across Table 4.2, the coefficients for the models do not change in ways very differently from the previous pooled OLS and 2SLS models. The fixed intercept remains relatively stable with a small standard error. The average temperature has a consistently positive effect on water consumption, and average precipitation has a negative effect (except in model RE-1T, where it is slightly positive, +0.11). An increase in sewer and fixed monthly costs both reduce water consumption, as do the indicators for 1992 and 2001, as can be observed in the second stage fits for models RE-2A and RE-2T. Most importantly, as in the transition from the pooled OLS to the pooled 2SLS model, implementing the Terza IV method in the 2S models flips the sign on the marginal price coefficient from positive to negative for both groupings. The coefficient of marginal price is -1.17 for RE-2A model, and -1.48 for the RE-2T model. The magnitude of the difference variable is almost negligible (0.04 and 0.08, respectively).

Comparing the two models fits requires several different metrics, because of the different number of random variables included in each model. For example, the models grouped by account all have significantly higher log likelihoods (a value of 76,387 for RE-1A and 62,193 for RE-2A) than the models grouped by tract (43,200 for RE-1T and 10,981 for RE-2T), but it is necessary to account for the significantly higher number of parameters that go into the models grouped by account. Examining the AIC values shows that even though the account models use significantly more parameters, they also result in significantly better (i.e., more negative) readings of the AIC criterion.

Similarly, the residual variance in the account-level models decreases. In the second-stage, model RE-2A has a residual variance of 2.51, while the account level model has a residual variance of 5.20. Although the within-group variance for the tract models is smaller (0.51, versus 3.06 for the account level model), this is likely a consequence of many more observations within each tract, since there are approximately 100 accounts within each tract. This effect can also be observed in Figure 4.2, in which the spread and standard errors of the varying intercepts decrease significantly. Increasing the number of groupings in the models with accounts significantly increases the errors, simply because there are less observations for each grouping.

However, Figure 4.3 also shows another relative trade-off between the models grouped by account and tract. When grouped by account, the plot of actual versus predicted values shows that these RE models are much more effective than previous OLS and 2SLS models at predicting higher-end values of water consumption. When grouped by tract, the averaging within each tract group continues to constrain predicted observations to a relatively tight band around the low-end of consumption. This is a particularly important problem for our policy analysis: because the policy of interest only affects high water users, we would like better predictions at high water consumption levels.

CRC models The main change between the RE and CRC models is that the coefficients are allow to vary by group: in the first stage models, the coefficients of the individual block prices are allowed to vary by account and tract group, respectively; in the second stage models, the coefficients of the predicted marginal price and difference variable instruments are allowed to vary by account and tract group. The intercepts are still allowed to vary by account and tract group, as in the RE models.

Fixed parameter estimates from the CRC models are shown in Table 4.3. The CRC models are grouped similarly to the RE models. For the intercept and slopes indicated by an asterisk (*), the values reported in this table are the means, and the random effects are drawn from a normal distribution with mean zero. The variances for the intercepts are shown in Figure 4.4, the variance of the coefficients are shown in Figures 4.4 – 4.6, and all the variances are summarized in Table 4.4.

The coefficient of marginal price and difference variable are the most important result from the CRC models. In Table 4.3, the mean coefficient for marginal price is -1.52 for the CRC-2A model and -1.74 for the CRC-2T model. Figures 4.4 – 4.6 also reveal important effects of the groupings. For the varying intercepts shown in Figure 4.4, the account groupings have significantly larger spread and standard error, as in the RE model. For the coefficient of marginal price and the difference variable, as shown in Figures 4.5 and 4.6, there is also a much larger spread of values, as well as the predicted standard errors, in the coefficients grouped by account over those grouped by tract.

The overall fit of the models, as shown in Figure 4.7, shows the tradeoffs made in the prediction process. Although the parameters obtained in the account-level model CRC-2A are much noisier, the plot of predicted values versus actual values shows a much more balanced fit in the upper ranges. In the tract-level model CRC-2T, the predicted values remain quite low in the upper ranges, and consistently under-predict water consumption.

Comparison with literature The main prediction of interest from the RE and CRC models is the average treatment effect of a change in marginal prices, or mean price elasticity. As discussed above, because the coefficients are described by statistical distributions, we cannot interpret the individual group coefficient values, so we examine the mean coefficients for marginal prices which are calculated as fixed parameters. The elasticity ϵ is calculated for the linear form as

$$\epsilon = \frac{b_{\hat{p}} p}{x}$$

where p and x are the price and quantity describe at a particular point on the demand curve. Using the average observed price and water consumption in each tier shown in Table 4.5, and correcting for the previous normalizations (multiply by standard deviation of 1.33, and divide by 100), we obtain the elasticities shown in Table 4.6. These are consistent with the lower-end values found by Dalhuisen et al. (2003) and Espey et al. (1997) in their meta-analyses of water demand studies, as discussed in Chapter 2.

4.6 Discussion

This chapter has introduced and applied the CRC model in modeling water demand. The CRC model is able to eliminate simultaneity bias as observed in literature, and as is demonstrated in the pooled OLS model. Compared to the pooled 2SLS model and RE models, which have been used in the literature before, the CRC model is able to allow flexibility and variation in the estimated coefficients of the marginal price and the difference variable. The inherent flexibility of the CRC model allows the fitted models to achieve higher AIC scores than the other model types, despite penalizing for additional parameters.

However, the inherent flexibility of the CRC model is also difficult to interpret. Since the coefficients obtained are described by parametrized statistical distributions, it is impossible to interpret the coefficient for any given individual or grouping. In this chapter, we therefore have to settle for calculating the average treatment effect, or expected value of the desired parameter (in this case, the coefficient of marginal price) within the finite group population. The elasticities that were calculated from the CRC model are in the range of values observed previously in the literature, but do not tell us anything about the individuals or groups. There are two alternatives to the CRC model that could be used to gain a better sense of individual- or group-level effects with respect to the marginal price.

First are the classical regression alternatives. Either a least squares dummy variable (LSDV) or 'fixed effects' model could be used to estimate individual- or group-level effects. A dummy variable could be used to indicate each grouping, and then interacted with the coefficient of marginal price to get an interaction coefficient that represents the coefficient of marginal price with respect to each individual or group, but in general multicollinearity between group-level indicators and predictors is a concern. There are also several weaknesses of the classical approaches: incorporating both group-level indicators and group-level predictors ignores possible interactions between individual- and group-level predictors; it is difficult to account for intrinsic group-level variation; and it is difficult to get reasonable estimates for small sample sizes, which is possible with accounts that have varying numbers of observations.

The second alternative is to develop a multilevel model, in which the variation in the intercepts and coefficients is explained in terms of other predictors at the group-level. Because of this dataset can be joined to other datasets that contain predictors related to the individuals and groups – and because this is a relatively flexible modeling technique – this is the approach that is described in the next chapter.

4.7 Tables & Figures

	Pooled OLS	Pooled 2SLS
Fixed intercept	26.93	26.93
	(0.05)	(0.06)
Average temp	1.88	3.14
	(0.05)	(0.11)
Average precip	0.11	-0.69
	(0.06)	(0.09)
Sewer	-3.54	-2.52
	(0.11)	(0.18)
Fixed cost	-7.26	0.30
	(0.11)	(0.20)
Indicator 1992	-0.37	-2.71
	(0.05)	(0.08)
Indicator 2001	-2.59	0.08
	(0.05)	(0.11)
Marginal price	24.35	-2.93
	(0.19)	(0.85)
Difference variable	10.15	4.34
	(0.18)	(1.92)
Residual se	19.71	28.05
R^2	0.3359	
Total observations	190,236	190,236
DOF	190,227	190,227

Table 4.1: OLS and 2SLS Results. All data is center-normalized, multiplied by 100 for ease of interpretation, and standard errors appear in parentheses (). As expected, applying the 2SLS model inverts the sign of the marginal price coefficient, and greatly reduces the magnitude of the both the marginal price and difference variable instruments. All other variables show the expected signs, except for the monthly fixed costs.

	RE model by account			RE model by tract		
	RE-1A	RE-2A		RE-1T	RE-2T	
	ObsMP	1S IV	2S IV	ObsMP	1S IV	2S IV
Intercept *	26.72	26.58	26.85	27.06	27.11	27.49
	(0.15)	(0.19)	(0.19)	(0.45)	(0.70)	(0.73)
Average temp	2.19	2.42	2.63	1.88	2.42	2.58
	(0.04)	(0.05)	(0.05)	(0.05)	(0.07)	(0.06)
Average precip	-0.10	-0.56	-0.39	0.11	-0.54	-0.39
	(0.04)	(0.05)	(0.05)	(0.05)	(0.07)	(0.06)
Sewer	-3.11	-2.20	-2.32	-3.53	-2.21	-2.27
	(0.09)	(0.10)	(0.09)	(0.11)	(0.13)	(0.13)
Fixed cost	-4.54	-0.77	-0.56	-7.10	-0.62	-0.22
	(0.09)	(0.14)	(0.10)	(0.11)	(0.19)	(0.14)
Indicator 1992	-1.22	-2.54	-2.48	-0.39	-2.58	-2.65
	(0.04)	(0.05)	(0.04)	(0.05)	(0.07)	(0.06)
Indicator 2001	-1.78	-0.03	-0.55	-2.55	-0.13	-0.59
	(0.04)	(0.07)	(0.04)	(0.05)	(0.10)	(0.06)
Price, block 1		1.59			1.66	
		(0.25)			(0.35)	
Price, block 2		-1.52			-1.68	
		(0.12)			(0.17)	
Price, block 3		-1.50			-1.26	
		(0.16)			(0.22)	
Marginal price	13.02		-1.17	23.53		-1.48
	(0.17)		(0.08)	(0.18)		(0.11)
Diff. variable	5.17		0.04	10.18		0.08
	(0.15)		(0.00)	(0.17)		(0.01)
Within-group var	1.87	2.93	3.06	0.20	0.47	0.51
Residual var	2.20	2.52	2.51	3.71	5.20	5.20
Log likelihood	76,387	62,278	62,193	43,200	10,987	10,981
AIC	(152,752)	(124,532)	(124,364)	(86,377)	(21,950)	(21,940)
BIC	(152,640)	(124,410)	(124,252)	(86,266)	(21,828)	(21,828)
Observations (N)	190,236	190,236	190,236	190,236	190,236	190,236
Groups (J)	8,741	8,741	8,741	97	97	97

Table 4.2: Fixed Effect Estimates for the RE Model Grouped by Account and Tract. All data is center-normalized, multiplied by 100 for ease of interpretation, and standard errors appear in parentheses (). Only fixed effects are reported, but the asterisk (*) indicates that this is the mean value of the intercept. Plots and histograms of the varying intercepts by account and tract appear in Figure 4.2, and are summarized in Table 4.4.

	CRC model by account			CRC model by tract		
	CRC-1A	CRC-2A		CRC-1T	CRC-2T	
	ObsMP	1S IV	2S IV	ObsMP	1S IV	2S IV
Intercept *	30.35	26.71	25.54	43.76	27.19	29.17
	(0.17)	(0.19)	(0.19)	(0.95)	(0.70)	(0.72)
Average temp	2.02	2.36	2.42	1.81	2.36	2.64
	(0.04)	(0.05)	(0.04)	(0.06)	(0.07)	(0.07)
Average precip	-0.16	-0.59	-0.38	-0.34	-0.55	-0.43
	(0.04)	(0.04)	(0.04)	(0.06)	(0.07)	(0.06)
Sewer	-2.57	-2.20	-2.13	-1.44	-2.27	-2.53
	(0.07)	(0.09)	(0.09)	(0.12)	(0.13)	(0.14)
Fixed cost	-2.72	-0.94	-0.28	1.61	-0.74	-0.51
	(0.08)	(0.12)	(0.09)	(0.15)	(0.19)	(0.16)
Indicator 1992	-1.21	-2.42	-2.35	-0.06	-2.50	-2.78
	(0.04)	(0.05)	(0.04)	(0.05)	(0.08)	(0.07)
Indicator 2001	-1.00	0.04	-0.63	-0.32	-0.09	-0.65
	(0.03)	(0.07)	(0.04)	(0.05)	(0.10)	(0.06)
Price, block 1 *		2.08			2.31	
		(0.26)			(0.46)	
Price, block 2 *		-1.60			-2.10	
		(0.15)			(0.37)	
Price, block 3 *		-1.72			-1.45	
		(0.15)			(0.23)	
Marginal price *	6.37		-1.52	5.81		-1.74
	(0.24)		(0.10)	(0.42)		(0.21)
Diff. variable *	-17.45		-0.33	-91.54		0.56
	(0.46)		(0.01)	(2.64)		(0.13)
Within-group var	2.10	2.97	2.89	0.85	0.47	0.42
Residual var	1.46	1.90	2.05	4.13	5.16	5.16
Log likelihood	97,944	72,035	71,923	32,745	11,461	11,448
AIC	(195,855)	(144,028)	(143,813)	(65,458)	(22,881)	(22,864)
BIC	(195,693)	(143,815)	(143,651)	(65,295)	(22,668)	(22,702)
Observations (N)	190,236	190,236	190,236	190,236	190,236	190,236
Groups (J)	8,741	8,741	8,741	97	97	97

Table 4.3: Fixed Effect Estimates for the CRC Model Grouped by Account and Tract. All data is center-normalized, multiplied by 100 for ease of interpretation, and standard errors appear in parentheses (). Only fixed effects are reported, but the asterisk (*) indicates that this is the mean value of the intercepts and coefficients. Plots and histograms of selected random effects appear in Figures 4.4–4.6 and are summarized in Table 4.4.

	RE model by account			RE model by tract			CRC model by account			CRC model by tract		
	RE-1A	RE-2A		RE-1T	RE-2T		CRC-1A	CRC-2A		CRC-1T	CRC-2T	
	Obsmp	1S IV	2S IV	Obsmp	1S IV	2S IV	Obsmp	1S IV	2S IV	Obsmp	1S IV	2S IV
Intercept	13.26	16.73	17.11	4.44	6.85	7.19	13.94	16.76	16.54	9.24	6.90	6.32
	(3.34)	(3.61)	(3.61)	(0.44)	(0.52)	(0.52)	(3.88)	(3.71)	(3.83)	(0.85)	(0.53)	(1.63)
P1								7.84			2.61	
								(7.06)			(1.21)	
P2								6.11			2.77	
								(6.09)			(1.38)	
P3								2.37			0.67	
								(2.44)			(0.12)	
MP							13.45		4.84	2.53		1.66
							(7.11)		(3.11)	(1.09)		(0.67)
DV							17.13		0.11	23.91		0.92
							(15.78)		(0.31)	(4.08)		(0.41)

Table 4.4: Standard Deviation of Random effects for RE and CRC models. All random variables for varying intercepts and coefficients are centered at zero, with the population-level means for the various reported random variables reported in Table 4.2 and 4.3. This table reports the standard deviation of the random variables, and the *mean* standard error, which is only provided to give an indication of the proportional errors in the random variable estimates.

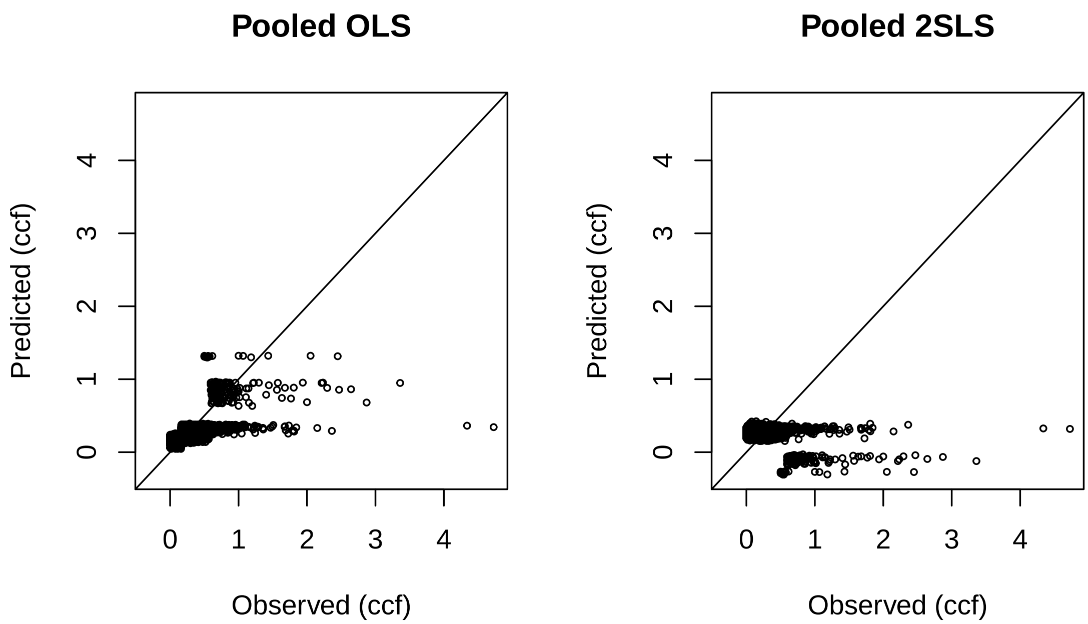

Figure 4.1: Predictions from Pooled OLS, 2SLS Versus Observed Data. From this plot we can see that the OLS model fit to the data has a clear structure of horizontal lines, which reflect the discrete (step) changes in marginal price over time. Since the higher-level observed values sit below the diagonal line, this model consistently under-predicts higher levels of water consumption. In the 2SLS model, the second-stage fit achieves a better fit by removing high predictions, but still consistently underpredicts outliers in actual consumption.

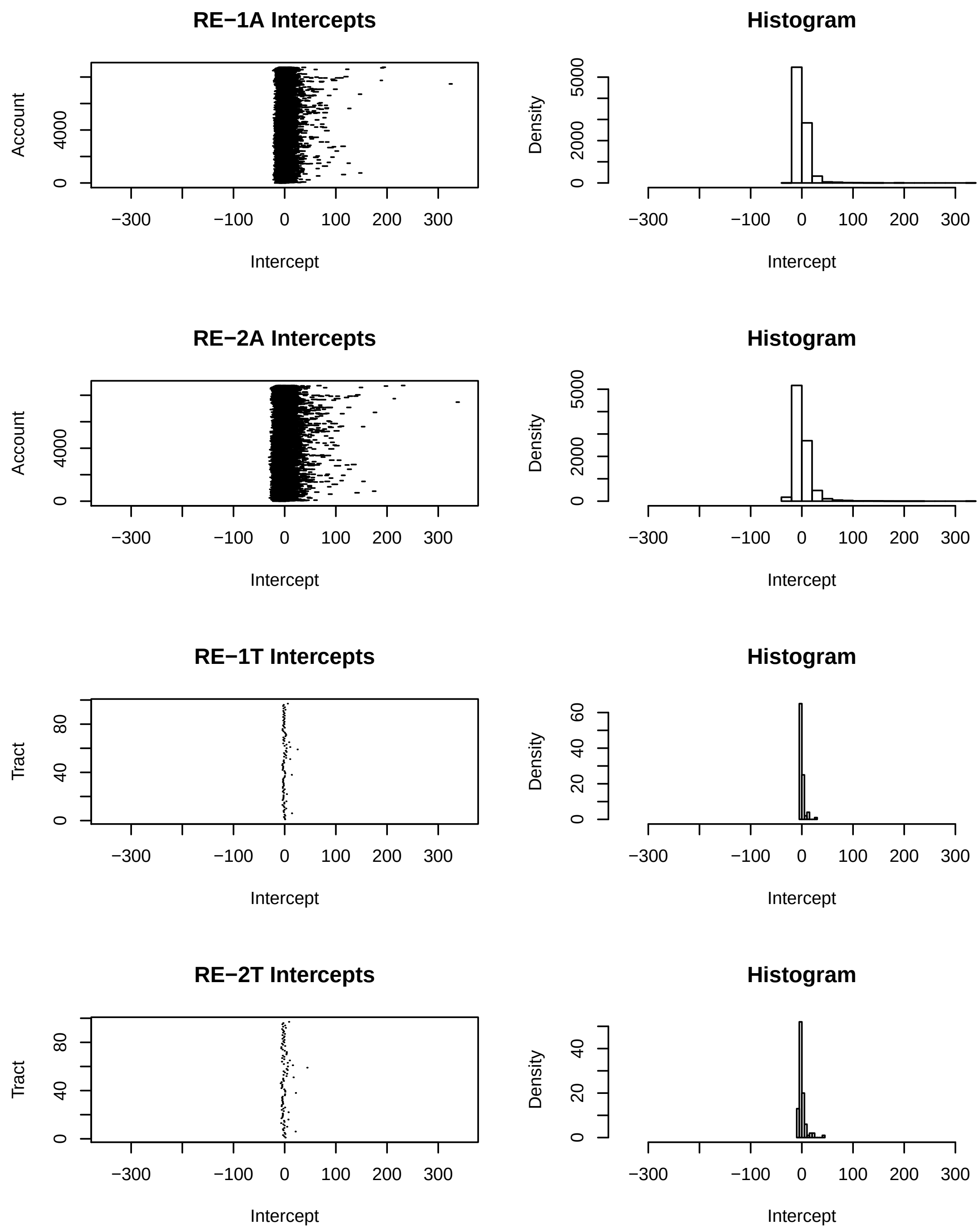

Figure 4.2: Estimate of Varying Intercepts from Random Effect Models. As can be seen from the account-level intercepts and histograms, there is considerable variation between accounts. Tract-level intercepts average out this variation to much lower values and tighter distributions.

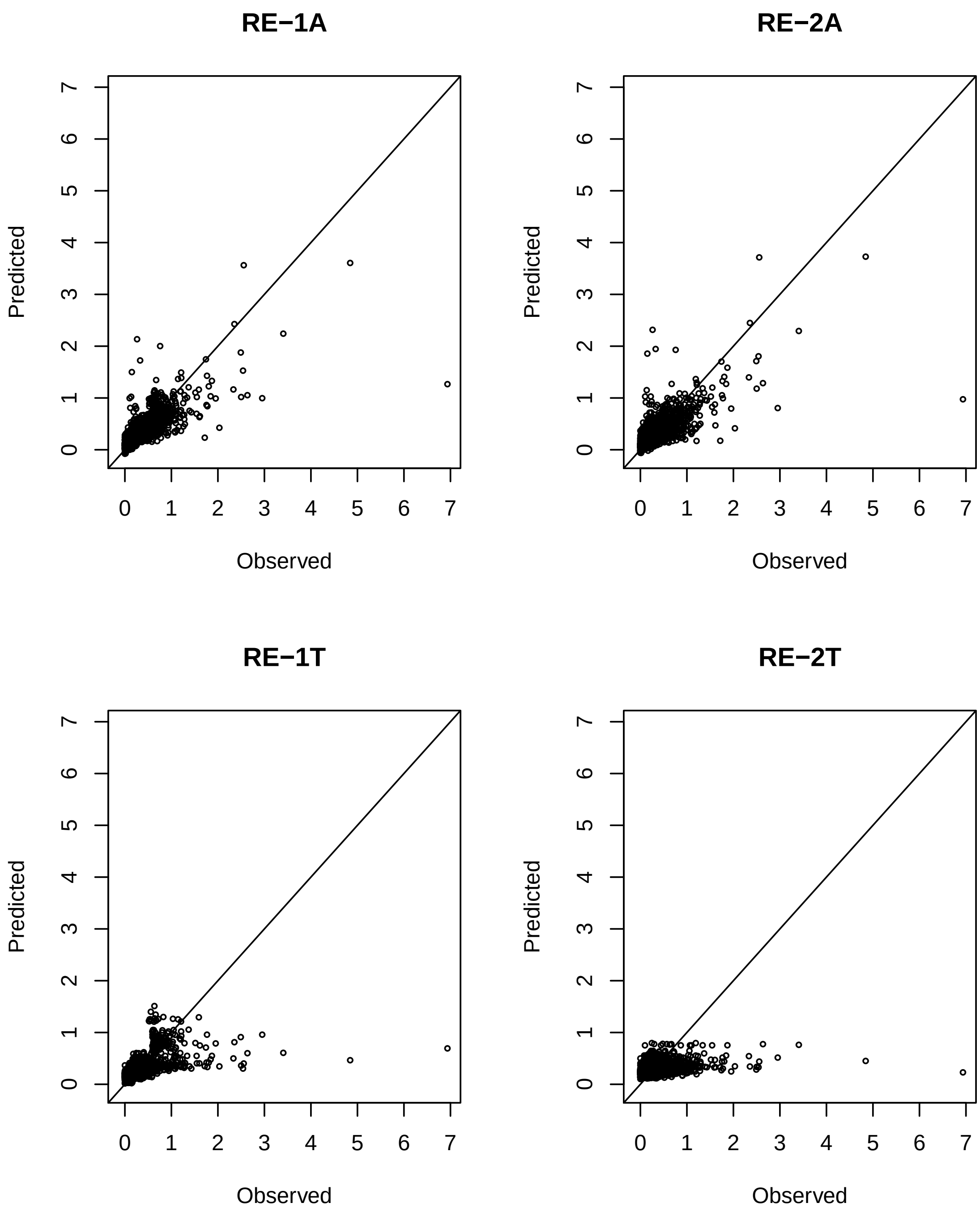

Figure 4.3: Predictions from Varying Intercept Models Versus Observed Data. Varying intercepts by account and tract allows models to fit higher-level water consumption observations. Error bars are not included since there is no classical interpretation of the error in the random variable intercepts.

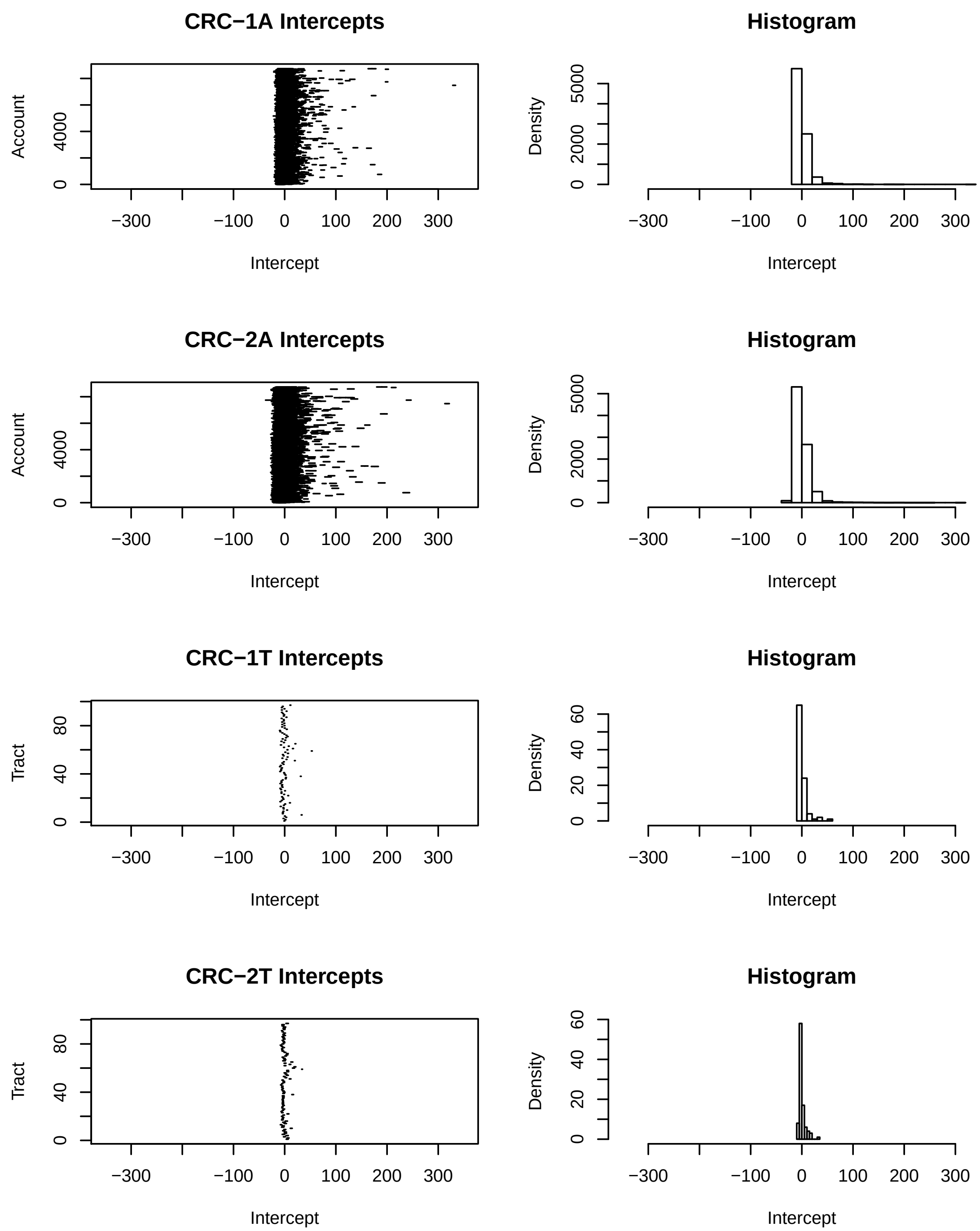

Figure 4.4: Estimates for Random Intercepts from CRC Models

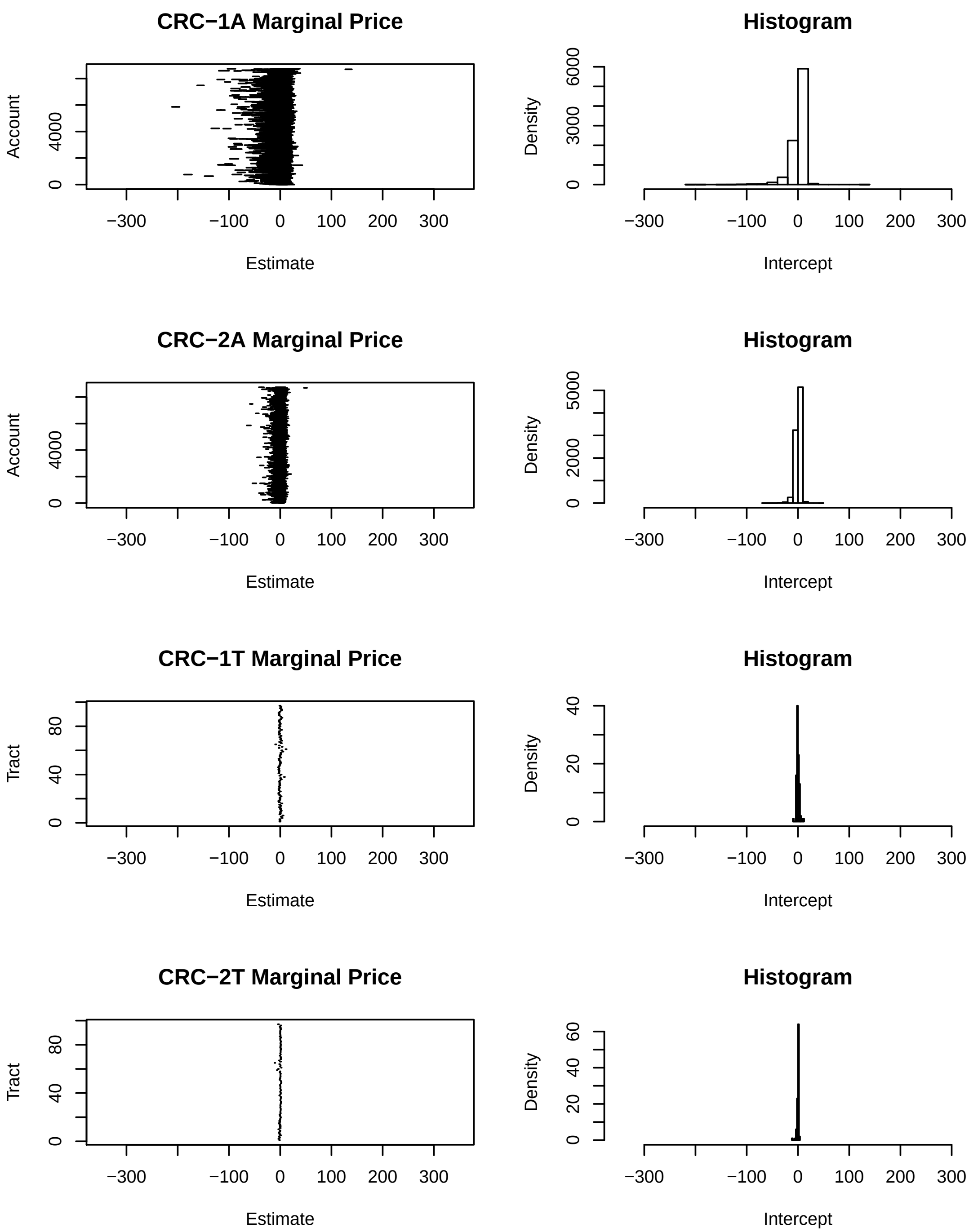

Figure 4.5: Estimates for Marginal Price Coefficient from CRC Models

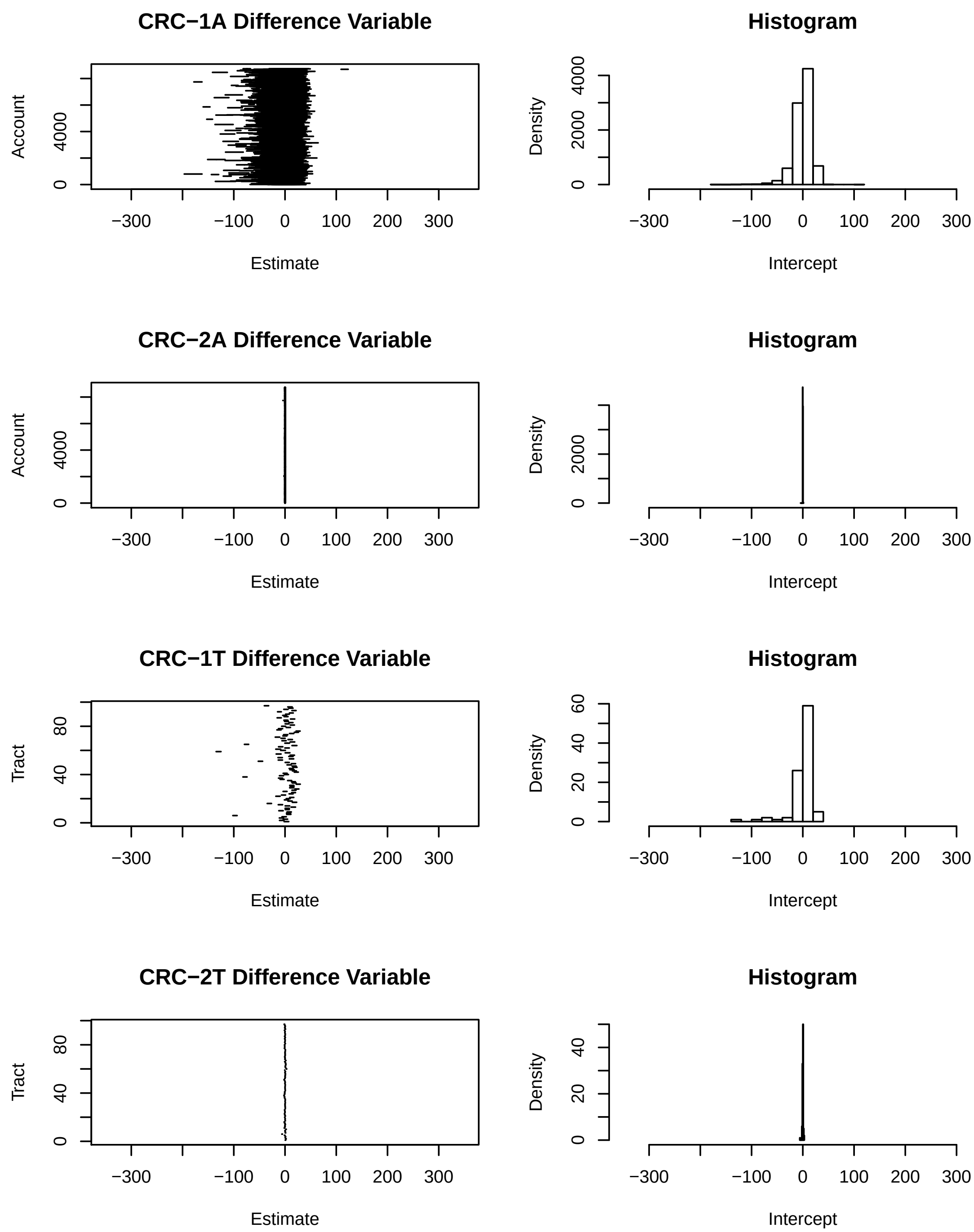

Figure 4.6: Estimates for Difference Variable Coefficient from CRC Models

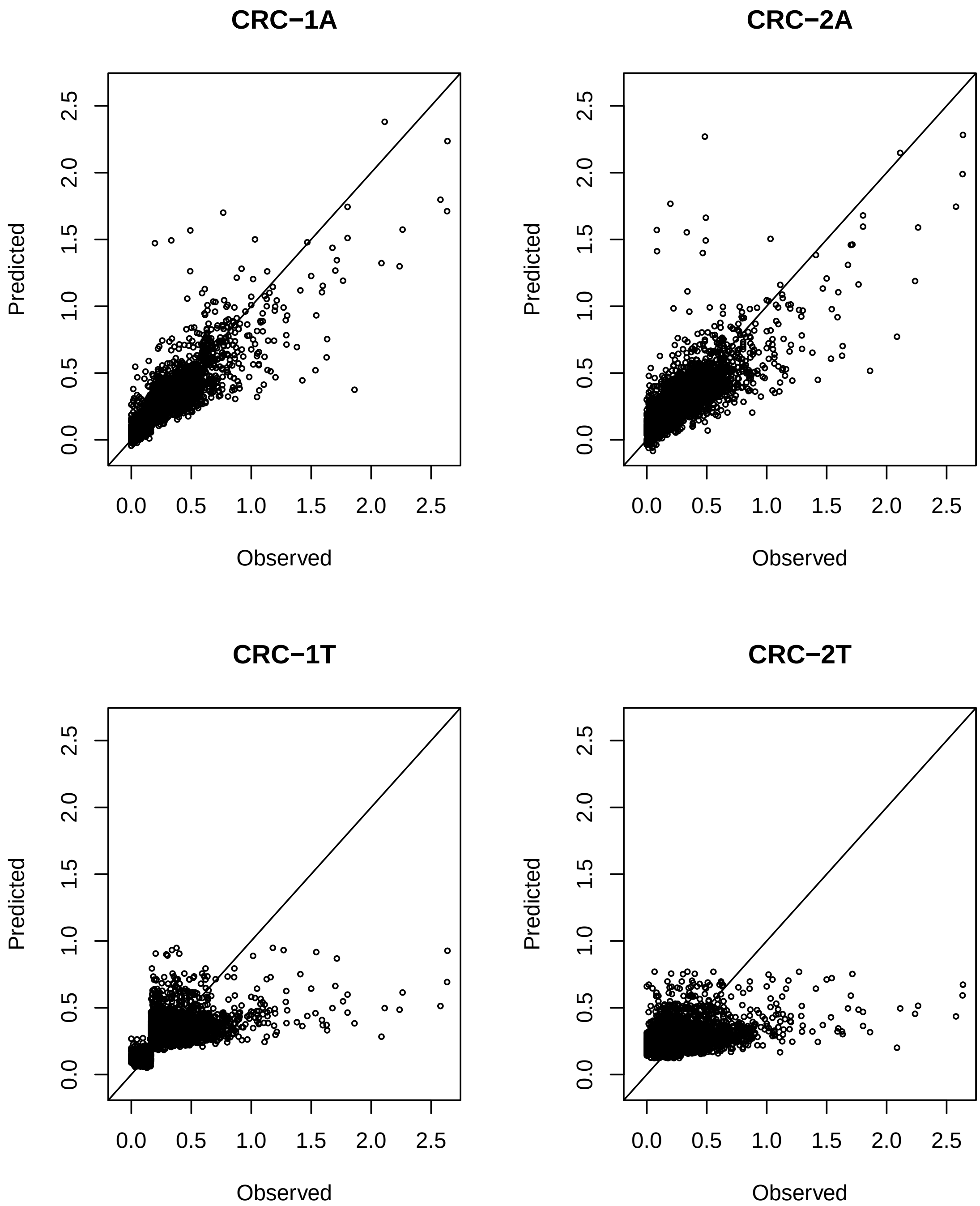

Figure 4.7: Predictions from CRC Models Versus Observed Data. Varying intercepts and coefficients by account and tract allows models to fit higher-level water consumption observations.

Tier	Mean P ($)	Mean Q (ccf)	Ratio
All tiers	3.15	0.27	11.63
Tier 1	2.48	0.11	22.58
Tier 2	3.15	0.32	9.79
Tier 3	9.75	0.80	12.16

Table 4.5: Average Price and Quantities By Tier

Model	MP coef	All tiers	Tier 1	Tier 2	Tier 3
RE-2A	-1.17	-0.18	-0.35	-0.15	-0.19
RE-2T	-1.48	-0.23	-0.44	-0.19	-0.24
CRC-2A	-1.52	-0.24	-0.46	-0.20	-0.25
CRC-2T	-1.74	-0.27	-0.52	-0.23	-0.28

Table 4.6: Calculated Elasticities for RE and CRC Models.

Chapter 5

HIERARCHICAL LINEAR MODEL OF WATER DEMAND

5.1 Introduction

The previous chapter introduced the use of correlated random coefficients (CRC) for the modeling of water demand, in order to explore how heterogeneous consumers react to pricing policies. The mean price elasticity – generally known as the average treatment effect – for heterogeneous consumers was found from the CRC model, and compared with results from the more familiar OLS, 2SLS and random effect (RE) models.

As the previous chapter concluded, however, though various estimated parameters for statistical distributions are estimated based on conditional parameters, in all of these classical models it is not possible to assess the average treatment effect for any *particular* cross-sectional unit. That is, if we use the CRC mathematical model as previously described:

$$\begin{aligned} Q_i &\sim \mathrm{N}(a_{j[i]} + \beta_i x_i + b_{j[i]}\hat{w}_i, \sigma^2) \\ \hat{w}_i &\sim \mathrm{N}(a^*_{j[i]} + \beta^*_i x_i + b^*_{j[i]} p_i, \sigma^2) \end{aligned}$$

with the hyperparameters

$$\begin{aligned} a_{j[i]} &\sim \mathrm{N}(\mu_a, \sigma_a^2) \\ b_{j[i]} &\sim \mathrm{N}(\mu_b, \sigma_b^2) \end{aligned}$$

then we can only calculate the price elasticity using the expectation across the distribution of the hyperparameter, that is,

$$\text{Price elasticity} \propto \mathrm{E}(b) = \mathrm{E}(b_{j[i]}) = \mu_b$$

In order to describe heterogeneity among individuals and groups in a meaningful way, we therefore need to be able to show how the random parameters are related to the inherent characteristics of those individuals and groups. This chapter explores the use of hierarchical linear models (HLM) in water demand, also known as multilevel models. As Gelman (2006) writes, "multilevel modeling is an increasingly popular analysis to modeling hierarchically structured data, outperforming classical regression in predictive accuracy. This is no surprise, given that multilevel modeling includes least squares regression as a special case".

HLM models introduce a second level of regression analysis, in which the group-level random parameters themselves can be related to group-level predictors. This second level is called the *group-level model*, in addition to the fit to the actual observed data, which is referred to as the *data-level model*. HLM analysis is similar to two-stage least squares regression, where group-level parameters from the first stage are then regressed again on other predictors. Or, put mathematically again, using the same data-level model, we now introduce specific distributions and parameters for each group, where,

$$\begin{aligned} a_{j[i]} &\sim \mathrm{N}(\alpha_{j0} + \alpha_j u_j, \sigma_a^2) \\ b_{j[i]} &\sim \mathrm{N}(\gamma_{j0} + \gamma_j u_j, \sigma_b^2) \end{aligned}$$

We can now calculate the price elasticity using the expectation for each group j, that is,

$$\text{Price elasticity} \propto \mathrm{E}(b) = \mathrm{E}(b_{j[i]}) = \gamma_{j0} + \gamma_j u_j$$

so we can now say something meaningful for each group.

HLM models have several key practical advantages over classical regression models, including those explored in Chapter 4. It remains possible to allow for systematic unobserved heterogeneity among groups, as in random parameter models. In addition, it is further possible to interact predictors at the data- and group-levels, in order to take into account interactions of a group with individual covariates. A further practical advantage of HLM models is that it is possible to join together and use disparate datasets by matching on groups, such as partitioning on geography or time periods. For example, in studying water demand, it is often difficult to obtain income information in the context of a water consumption survey. This chapter will also demonstrate how other datasets, such as property tax assessments and Census data, can be applied to provide proxy or aggregate information about income. Another theme explored in the previous chapter was assessing the efficiency of models grouped in different ways, and HLM models allow characterization of explained variance at each of the multiple levels.

Most importantly for the purposes of this chapter, HLM models are extremely useful in assessing prediction results. In HLM models, the group-level regression against group-specific predictors takes into account the estimation uncertainty of the group-level parameters themselves. Uncertainty at the group-level is introduced from a number of sources, including:

- from unbalanced samples, in which groups have different sample sizes;
- from individual-level uncertainty and group-level means, which are sometimes referred to as 'direct' and 'contextual' effects, respectively;
- variations in fit, where predictors may explain one group much better than in another.

Propagation of error through the model ultimately allows us to assess the error in the model predictions.

Finally, Bayesian techniques allow us not only to estimate HLM models, but are consistent with the interpretation of errors. As mentioned above, in the classical or frequentist perspective there are no errors in the observation of a random variable, even when these are obtained from a statistical distribution with estimated parameters; and it does not necessarily make sense to add these errors to the estimated standard errors of the fixed effects. In the Bayesian perspective, there is no such inconsistency, because both the 'fixed' and 'random' effects are intrinsically uncertain. Errors in the estimated parameters of a statistical distribution can be propagated to the observations from this distribution, analytically or using simulation, and then added directly to the errors in other estimated parameters. This chapter will use simulation to obtain errors in the predictions which are necessary for model assessment.

This chapter will further build on the CRC models introduced in Chapter 4, using HLM models in order to demonstrate some of these advantages, and to obtain predictions for use in the policy analysis in Chapter 6.

5.2 Model Specification

The HLM model used in this chapter specifies is an extension of the CRC model used in the previous chapter. Restating equation 4.18, the CRC model adds the correlated random coefficients $\mathbf{B_{j[i]}}$ varying by group:

$$
\begin{aligned}
Q_{ijt} &\sim \mathrm{N}(\boldsymbol{\beta_i X_i} + \mathbf{B_{j[i]}}\hat{\mathbf{w}}_{ijt}, \sigma^2) \\
\mathbf{B}_{j[i]} &\sim \mathrm{MVN}(\mu_{\mathbf{j}}, \mathbf{\Sigma_j}^2)
\end{aligned}
$$

In the HLM model, we can change the second line so the parameters of the varying coefficients themselves depend on group-level predictors:

$$
Q_{ijt} \sim \mathrm{N}(\boldsymbol{\beta_i X_i} + \mathbf{B}_{j[i]}\hat{\mathbf{w}}_{ijt}, \sigma^2) \tag{5.1}
$$

$$
\mathbf{B}_{j[i]} \sim \mathrm{MVN}(\mathbf{\Gamma_j u_j}, \mathbf{\Sigma_j}^2) \tag{5.2}
$$

where the correlated random coefficients $\mathbf{B}_{j[i]}$ are drawn from a multivariate normal distribution with mean conditional on $\mathbf{\Gamma_j}$, the matrix of group-level coefficients, matrix multiplied times the vector of group-level predictors $\mathbf{u_j}$. The variance structure is described by the covariance matrix $\mathbf{\Sigma_j}$. As in the previous chapter and in specifications with less random effects, $\boldsymbol{\beta}$ is the vector of coefficients for the individual observation-level predictors $\boldsymbol{X_i}$.

It may further clarify the structure of equation 5.2 to write it out element-by-element. In the matrix of correlated random coefficients $\mathbf{B}_{j[i]}$, we have $J \times K$ correlated random coefficients $b_{j[i]k}$, one for each of the J groups and each of the K group-level covariates (in our application, usually treatments or instruments). In turn, each of these elements $b_{j[i]k}$ is modeled as an observation from the statistical distribution parametrized on γ_{jkl} for L group-level predictors u_{jkl}:

$$
b_{jk} \sim \mathrm{N}(\gamma_{jkl} u_{jkl}, \sigma_k^2) \tag{5.3}
$$

where diagonal elements of the covariance matrix $\mathbf{\Sigma_j}$ are written as σ_k^2 (since covariances need to be expressed in matrix form).

5.3 Instrument Specification

The literature suggests numerous possible ways to specify the instruments in the first stage (see Chapter 2). This paper applies two methods appropriate to the detailed microdata in this study. The earlier IV method developed was by Wilder and Willenborg (1975), in which the observed marginal price and difference variable are regressed on the individual block prices:

$$\hat{P}_i \sim \mathrm{N}(\mathbf{a} + \boldsymbol{\beta_i X_i} + \mathbf{b}_p\mathbf{p}, \sigma_P^2) \tag{5.4}$$

$$\hat{D}_i \sim \mathrm{N}(\mathbf{a} + \boldsymbol{\beta_i X_i} + \mathbf{b}_p\mathbf{p}, \sigma_D^2) \tag{5.5}$$

where $\mathbf{p}$ is the vector of individual block prices, $\mathbf{b}_p$ is the vector of coefficients for each block price, and all other variables β_i can be included, though they are estimated specifically for each regression in the first stage.

The Wilder-Willenborg IV (WWIV) method has not been frequently used in the literature, probably due to the limited availability of detailed price information for individual accounts. By applying a linear regression to represent the nonlinear block rate price structure, some call the WWIV method an attempt to linearize or 'capture the spirit of' the price structure. Although this method has the disadvantage of predicting marginal prices and difference variables that may not actually exist between blocks, the WWIV method has the interesting advantage of maintaining a direct linear relationship between the individual block prices; the marginal price and difference variable instruments; and the observed outcome.

This direct linear relationship is also particularly useful to understand the effect of changing the third block price on the observed consumption. One persistent question in the literature has been an understanding of what price elasticity actually means in a nonlinear block rate price structure: Is it the price elasticity with respect to the average of the individual block prices? With respect to all of the block prices? Or, with respect to the observed block price? As Olmstead et al. (2007) note, the literature has never quite clarified this topic, and they do not seem to offer an answer themselves.

In the WWIV method, the direct linear relationship of the individual block prices to the observed marginal price in the first stage, and of the marginal price instrument to the outcome in the second stage, allows us to calculate directly the local average treatment effect (LATE), a ratio which captures the change in outcome observed as a result of changing the underlying predictors of the instrument itself:

$$\text{LATE} = \frac{\text{2S coef of Q wrt to instrument}}{\text{1S coef of instrument wrt to predictor}} \tag{5.6}$$

Expressed this way, the LATE can be used in calculating the quantity change with respect to price changes in *only* the third tier. In the causal interpretation of IV methods, this only applies for those consumers who are likely to switch, or could be induced to switch by price changes. Gelman and Hill (2007, chapter 10) and Angrist and Pischke (2009, chapter 4) discuss some of the restrictions on causal interpretation of the instrumental variables.

Interpreting the effect of changing the third tier price on the observed consumption as a LATE is a novel addition to the water demand literature.

As discussed in Chapter 4, Terza (1986) offered a useful correction to the tendency of the Wilder-Willenborg IV method to predict observed marginal prices and difference variables that do not actually exist. Instead of using a predicted (and linearized) marginal price and difference variable instrument as a predictor in the second stage, Terza instead predicted the quantity of consumption based on the individual block prices, and then assigned a marginal price and difference variable based on the actual rate schedule. This method has the advantage of first predicting the quantity on the observable predictors related to changes in the price schedule, and then assigning plausible instruments for the second stage. However, the direct linear relationship of the predictors of the instruments, and of the instruments to the outcome, is lost.

Both instruments will be employed in the following analysis, and the results will be compared and discussed.

5.4 Estimation Theory & Methods

This section will present a very brief summary of the theory behind estimation of a multilevel model such as equations 5.1 and 5.2, based on the more comprehensive descriptions of the theory of Bayesian estimation in Lee (2004), Gelman et al. (2004), and Gelman and Hill (2007, chapter 18). This brief description will start with conditional probability and Bayes' Rule to demonstrate the essence of Gibbs sampling, and to demonstrate how this is used in a multilevel model.

In Bayesian estimation we are generally interested in estimating a vector of k parameters $\boldsymbol{\theta} = (\theta_1 \ldots \theta_k)$ conditional on n observations $\boldsymbol{X} = (X_1 \ldots X_n)$, or

$$p(\boldsymbol{\theta}|\boldsymbol{X})$$

This original relationship of the vector of parameters to the original data is called the *posterior*, emphasizing the idea that this is an *a posteriori* relationship, that is, obtained by deduction or observation.

Making the initial assumption that the vector of observed data depends on the parameters in a known way, for example in a parametrized statistical distribution, then we can write this relationship as a pdf:

$$p(\boldsymbol{X}|\boldsymbol{\theta})$$

Inverting this relation and writing it as a function of $\boldsymbol{\theta}$ on X, this term is referred to as the *likelihood* l:

$$p(\boldsymbol{X}|\boldsymbol{\theta}) = l(\theta|\boldsymbol{X})$$

In order to link these two conditional probabilities, we need some *a priori* beliefs about the distribution of the parameters, which we express in terms of a pdf:

$$p(\boldsymbol{\theta})$$

which we refer to as the *prior distribution.* The basic properties of conditional probability are restated as Bayes' Rule to link these three elements together:

$$\begin{aligned} p(\boldsymbol{\theta}|\boldsymbol{X}) &\propto p(\boldsymbol{\theta})p(\boldsymbol{X}|\boldsymbol{\theta}) \\ &\propto p(\boldsymbol{\theta})l(\theta|\boldsymbol{X}) \end{aligned}$$

or, put simply,

$$\text{posterior} \propto \text{prior} \times \text{likelihood}$$

We can then apply Bayes' Rule sequentially to obtain estimates from mutliple independent observations. If and only if a set of observations Y is independent of X, then

$$p(\boldsymbol{X}, \boldsymbol{Y}|\boldsymbol{\theta}) = p(X|\boldsymbol{\theta})p(Y|\boldsymbol{\theta})$$

and trivially we can see that

$$l(\boldsymbol{\theta}|\boldsymbol{X}, \boldsymbol{Y}) = l(\boldsymbol{\theta}|\boldsymbol{X})l(\boldsymbol{\theta}|\boldsymbol{Y})$$

We can then restate the probability of $\boldsymbol{\theta}$ on the joint probability of X, Y as:

$$\begin{aligned} p(\boldsymbol{\theta}|\boldsymbol{X}, \boldsymbol{Y}) &\propto p(\boldsymbol{\theta})p(\boldsymbol{X}, \boldsymbol{Y}|\boldsymbol{\theta}) \\ &\propto p(\boldsymbol{\theta})l(\boldsymbol{\theta}|X)l(Y|\boldsymbol{\theta}) \\ &\propto p(\boldsymbol{\theta}|X)l(\boldsymbol{\theta}|Y) \end{aligned}$$

The posterior for $\boldsymbol{\theta}$ can therefore be found by treating the first term, $p(\boldsymbol{\theta}|\boldsymbol{X})$ as the prior for the likelihood $l(\boldsymbol{\theta}|\boldsymbol{Y})$. This can be conducted iteratively for any number of parameters and is independent of order.

Sequential estimation of likelihoods for $l(\boldsymbol{\theta}|\boldsymbol{Y})$ and $l(\boldsymbol{\theta}|\boldsymbol{X})$ is, in essence, what *Gibbs sampling* does. A Gibbs sampler is an iterative algorithm that partitions the space of parameter values, and given a set of initial values and prior distributions, iteratively estimates the parameters in sequential order. Each parameter is in turn drawn from prior distributions of the other parameters, until all of the parameter values approach some pre-defined level of convergence from multiple starting values.

Sequential estimation by Bayesian estimation is critical in multilevel modeling. In the multilevel model, the group-level model defines the prior distributions for the data-level model. That is, starting with initial values for the group-level coefficients $\mathbf{\Gamma}_j$ and variance matrix $\boldsymbol{\Sigma_j}$ in equation 5.2:

$$\mathbf{B}_{j[i]} \sim \mathrm{MVN}(\mathbf{\Gamma_j u_j}, \mathbf{\Sigma_j}^2)$$

we can then draw values for $\mathbf{B}_{j[i]}$. These are in turn combined with prior distributions for β_i and σ to serve as priors to estimate the data-level model in equation 5.1:

$$Q_{ijt} \sim \mathrm{N}(\beta_i X_i + \mathbf{B}_{j[i]} \hat{\mathbf{w}}_{ijt}, \sigma^2)$$

In the multilevel model, the data-level model (equation 5.1) is frequently referred to as the *likelihood* and the *group-level model* (equation 5.2) is frequently referred to as the *prior model*. Gibbs sampling repeats this sequence of prior and posterior calculations, alternating between the data- and group-levels, until the parameter values converge from multiple starting values, which yields the most likely dependence of the entire parameter set on the observed data.

The choice of prior distributions are therefore key assumptions necessary to estimate the multilevel model using Gibbs sampling. For the parameters that do not have any explicit dependence on observations, or about which is little is known, non-informative prior distributions are used so that as little prior information is put in the model as possible, and the Gibbs sampler then converges towards the likely values based on the data. In order to estimate the $J \times K$ different elements in the matrix $\mathbf{B}_{j[i]}$ in the group-level equation 5.2, the initial values of each element $\boldsymbol{\Gamma_j}$ are drawn from the normal distribution with a mean of zero and a variance of 10,000, and the initial values for the elements of the $K \times K$ covariance matrix are drawn from an inverse-Wishart distribution with $K + 1$ degrees of freedom. In order to estimate predicted values for the N observations of water consumption, the initial values for β_i are also drawn from a normal distribution with mean zero and a variance of 10,000, and σ is drawn from a uniform distribution on the range (0,100). As in

Chapter 4, all of the observed data at the data- and group-levels was center-normalized to aid in convergence, as well as to aid in interpretation and comparison of results.

As in the previous chapter, we will test various groupings by individual households or accounts, as well as by census tract. The key issue therefore in assessing goodness of fit is how well the various groupings and group-level predictors explain the observed variations.[1]

Computational limits played a much larger role in the Bayesian estimation and analysis, compared to the classical models presented in Chapter 4, so it is worth discussing here. Code for the Gibbs sampler written in the BUGS language appears in Appendix A. Gibbs sampling was performed using the JAGS (Just Another Gibbs Sampler) module developed by Plummer (2010), and accessed from a 64-bit build of the R language (version 2.10.1) using the R2jags package developed by Su and Yajima (2010). With approximately 10,000 data points, the code required approximately one week to execute on an Apple iMac with a 2.4 GHz Core 2 Duo processor and 4 GB of RAM. The discussion section at the end of

---

[1]Another good metric for assessing the explanatory power of multilevel models at various levels are the Bayesian measures of explained variance and pooling, developed by Gelman and Pardoe (2006). The fit of linear regression models for $i = 1 \dots n$ observations is often summarized by:

$$R^2 \quad = \quad 1 - \frac{\mathrm{V}_{i=1}^{n} \epsilon_i}{\mathrm{V}_{i=1}^{n} y_i} \tag{5.7}$$

where $\mathrm{V}_{i=1}^{n}$ is the finite-sample variance operator, $\mathrm{V}_{i=1}^{n} x_i = \frac{1}{n-1} \sum_{i=1}^{n} (x - \hat{x})^2$. Gelman and Pardoe (2006) note that in a multilevel model, variation comes from multiple levels. For each level m, the variation comes from

$$\theta_k^m = \mu_k^m + \epsilon_k^m \tag{5.8}$$

where $k = 1 \dots K^{(m)}$ is the number of observations in each level m, variations θ_k^m come from the linear predictors μ_k^m, and the random effects ϵ_k^m. At various levels of the model, these variations have different sources. At the data-level, θ_k^m come from the individual data points, and at the group-level, θ_k^m come from the modeled intercepts or coefficients. Gelman and Pardoe (2006) generalize the expression for R^2 into an analogous expression for each level,

$$R^2 \quad = \quad 1 - \frac{\mathrm{E}(\mathrm{V}_{i=1}^{n} \epsilon_k)}{\mathrm{E}(\mathrm{V}_{i=1}^{n} \theta_k)} \tag{5.9}$$

where E is the posterior mean. This measure of explained variance can be simulated for each level from the model with estimated parameters. A pooling factor can also be calculated for each level to summarize the extent to which variance of the errors ϵ_k is reduced by grouping, defined by Gelman and Pardoe (2006) as:

$$\lambda \quad = \quad 1 - \frac{\mathrm{V}_{i=1}^{n} \mathrm{E}(\epsilon_k)}{\mathrm{E}(\mathrm{V}_{i=1}^{n} \epsilon_k)} \tag{5.10}$$

The R^2 equivalent equation 5.9, and the pooling factors λ, have not been calculated for each grouped model; this remains for future work.

this chapter will consider various alternatives to speed up these calculations.

5.5 Results

The following discussion of the results from the various models is itself hierarchically ordered, first by groupings at the data-level, second by the instrumental variable method, and third by the group-level regression results. So, models grouped by account will be first discussed, including models using the observed marginal price and difference variable (i.e., no IVs used), the Terza IV, and the WWIV method, and the underlying group-level regression results discussed. Then, models grouped by tract and the associated results for the various IV methods will be discussed. Finally, key metrics and conclusions will be systematically explored, including the relative levels of explained variance and pooling factors in each model formulation, and the dependence of price elasticities on individual household- and census-tract level characteristics.

Models Grouped By Account For the models grouped by account, model A1 comes from a HLM fit using the observed marginal price and difference variable (i.e., no IVs used), models A2 and A3 represent the first and second stages, respectively, using the Terza IV; and models A4, A5, and A6 represent the two first stages and one second stage, respectively, of the WWIV method. All of the models are estimated using 10,489 billing observations grouped into 485 accounts.

Table 5.1 shows the data-level parameters from the account-level models, estimated using Bayesian inference. Numbers represent the mean simulated values from the posterior distribution, and numbers in parentheses () represent the sample standard deviation of the posterior simulated values, similar to standard errors. Plots and histograms for the data-level random effects appear in Figures 5.1-5.3. Group-level coefficients appear in Tables 5.2-5.7.

All of the fixed effects at the individual level appear to be broadly similar to the classical CRC models obtained in Chapter 4. The mean intercept value is ranges from 26.81 (model A1) to 19.72 (model A6). All of the coefficients show the expected signs, with an increase in average temperature causing an increase in consumption, and an increase in average precipitation, sewer costs, fixed costs, and indicators all causing a decrease in consumption. What is most striking about these model results are the relatively small posterior standard deviations compared to the classical estimates of standard error. Using only approximately 5% of the data in Chapter 4 gives significantly better estimates in the final model stages A1, A3, and A6.

Unfortunately, for the most important two variables relevant to pricing policy – the marginal price and difference variable coefficients – are decidedly inconclusive with much larger errors. In model A1, the marginal price coefficient (+8.79) is very large, but almost equal to the estimated variation (7.04). In model A3, the marginal price coefficient (-0.24) is almost insignificant compared to the estimated variation (7.67). In model A6, the marginal price coefficient (4.12) is also obscured by wider estimated variation (7.35). Figures 5.2 and 5.3 show the relatively wide spread around zero and do not clearly indicate if the coefficients for marginal price and difference variable are clearly positive or negative, though all the outliers for marginal price tend to be negative.

However, if we examine the group-level coefficients in Tables 5.2-5.7, we find that there are some interesting relationships between the estimated coefficients at the data-level and the group-level predictors. We will closely examine Table 5.2 first to demonstrate how to read this set of tables, and then comment more broadly on the more interesting IV models, particularly the second stage estimates for model A3 in Table 5.4 and model A6 in Table 5.7.

The group-level coefficients for model A1 are shown in Table 5.2; these estimates are of the parameters in equation 5.2. Most of the coefficients at the group-level are significant with relatively small standard deviations. In the first row, the data-level intercept value β_0 is strongly anchored by the group-level intercept γ_0 with a value of 31.06, which is larger than the coefficients γ with respect to lot size or house value, though both are significant. Only house age is relatively insignificant. In the second row of Table 5.2, the data-level coefficients for marginal price β_P actually has a positive mean γ_0 of +4.43 with a variation of 1.09. (This positive sign similarly motivated us to use IV methods in Chapter 4). However, the coefficients γ with respect to lot size and house value are strongly negative, meaning that a one standard deviation change in a larger and more expensive house would result in a significantly negative price elasticity. In the third row of Table 5.2, the group-level coefficient on difference variable has a strong negative mean (-21.43) with little variation (2.35), while the coefficients γ with respect to lot size and house value are much more weakly negative.

Now armed with a clearer explanation of how to read the group-level parameter tables, we turn to the second stage estimates for model A3 in Table 5.4 and model A6 in Table 5.7. The coefficient β with respect to marginal price appears in the second row of Table 5.4. The mean coefficient is strongly negative (-2.62) with a moderate standard deviation (+1.08), which will result in a significantly higher price elasticity than estimated in Chapter 4. Furthermore, because the coefficients γ with respect to lot size and house value are strongly negative (-2.26 and -3.88, with standard deviations of 0.83 and 0.85, respectively), a larger and more expensive house will have significantly higher negative price elasticities. This corresponds to our intuition that either larger houses or higher incomes are related to discretionary water use. In Table 5.7, the mean marginal price coefficient is actually positive (1.74) with a variation of 1.05, but the coefficients γ with respect to lot size and house value have very similar values as in model A3 (-2.30 and -3.86, with variations of 0.87 and 0.84, respectively).

Model predictions in Figure 5.4 show a relatively good clustering around the diagonal line, indicating the predicted values are relatively good compared to the actual observations at both low and high values of water consumption. Importantly, as explained in the introduction to this chapter, error bars can be put on the predicted values. Figure 5.5 shows that six randomly chosen accounts generally reflect the low average water usage, and the actual observations are well within the prediction error bars. Figure 5.6 shows that six randomly chosen high-consuming accounts. Here, although not all observations are within the prediction errors, the account-level model still manages to center the predictions relatively close to the actual observations, except with extreme outliers.

Models Grouped By Tract For the models grouped by census tract, model T1 comes from a HLM fit using the observed marginal price and difference variable (i.e., no IVs used), models T2 and T3 represent the first and second stages, respectively, using the Terza IV; and models T4, T5, and T6 represent the stages of the WWIV method. All of the models are estimated using 10,489 billing observations grouped into 97 census tracts.

Table 5.8 shows the data-level parameters from the account-level models, estimated using Bayesian inference. Numbers represent the mean simulated values from the posterior distribution, and numbers in parentheses () represent the standard deviation of the simulated values, similar to standard errors. Plots and histograms for the data-level random effects appear in Figures 5.7-5.9. Group-level coefficients appear in Tables 5.9-5.14.

Reading the results grouped by tract similarly to the previous results, we can see that the majority of the 'fixed' effects or control variable coefficients remain the same. The intercept falls in the same range (26.95 for model A1 ranging to 24.26 for model A6, with small variations of 1.83 and 2.88), and all of the covariates show the expected signs with relatively small variations. As in the models grouped by account, the HLM model is able to predict the coefficients for the observation-level predictors relatively accurately with small variations.

This time, when we examine the data-level IV coefficients, we see similar estimates to the CRC models in Chapter 4. In model T1, which uses the observed marginal price and difference variable, the marginal price coefficient is strongly positive (+22.33) with a small variation of 4.44, and the difference variable is 5.24 with a relatively large variation of 5.05. In model T3, as in the classical CRC models, the Terza IV method flips the sign of the marginal price coefficient (-2.11), though the standard deviation is relatively larger (4.55) as in the account model A3. Similarly, in model T6, the WWIV method flips the sign of the marginal price coefficient (-1.31), though again the standard deviation is larger (3.95).

However, when we examine the group-level parameter tables for the models grouped by tract (Tables 5.9-5.14), we see that the group-level coefficients have much larger errors than in the models grouped by account. Aggregate group-level predictors such as the average household size, median income, and median value could be expected to roughly to correspond to the predictors of lot size and house value for each account. However, the coefficients for all of the group-level predictors are considerably more noisy, with almost all of the group-level coefficients γ for the coefficients on marginal price and difference variable β dominated by their variations or simulation error. This has several possible causes. First, there are many less census tracts than accounts (97 versus 485), so less groups makes it actually more difficult to observe variation in the group means. Second, it is possible that aggregated group-level predictors are simply poor predictors of individual observations or of group means. Third, these could simply be the wrong predictors to compare with the account-level.

One immediate conclusion in comparing the data grouped by account and tract is that all of these possible causes indicate that tract-level models are likely to appear worse than account-level models as we scale up the amount of data. All of the advantages of the account-groupings will increase with the scale of the data, because as datasets become larger, although they will also contain corresponding larger numbers of accounts, the number of census tracts will not change. It may be possible to use other intermediate geographical groups, although other associated group-level predictors must also be found.

Elasticity Calculations The mean price elasticity can be calculated directly from the data-level marginal price coefficients presented in Tables 5.1 and 5.8, but the large variation in the estimated coefficients unfortunately do not result in any significant estimates of the mean price elasticity. Because the elasticity is linearly related to the data-level coefficients for marginal price, the elasticities for each individual account and tract will simply be rescaled versions of the plots and histograms of the marginal price coefficient in Figures 5.2 and 5.8. Unfortunately, these elasticities are not particularly informative because of the large number of positive elasticities. Elasticities plotted by tract appear in Figure 5.13. Average levels of consumption are shown in Figure 5.14.

Instead of using the data-level estimates, which contain significant variation, the mean price elasticity in this chapter can also be calculated directly from the γ coefficients obtained using the HLM analysis. We can get the price elasticity for an average household or tract by taking the intercept values for the β coefficient of marginal price, which correspond to the mean of the normally distributed values. Table 5.16 shows the price elasticities for a household that is average in all respects. Using the γ coefficients also allows us to develop a sensitivity analysis for the price elasticity to different household and tract characteristics. Tables 5.17 and 5.18 shows the calculated elasticities for an average household account and an average census tract, respectively, and with one standard deviation changes in their group-level characteristics.

LATE Calculation Similar to the elasticities, the LATE can be calculated for specific users and also using the group means using the results of the WWIV models. The LATE is usually calculated using the estimated coefficients from the first- and second-stages:

$$\text{LATE} \quad = \quad \frac{\text{2S coef of Q wrt MP}}{\text{1S coefficient of MP wrt P3}} \tag{5.11}$$

where MP is the marginal price instrument influenced by the individual block price P3. Since the second-stage models are calculated using two instruments, for both the marginal price and difference variable, and since all of the coefficients are center-normalized, I can calculate the LATE as a sum of the ratios:

$$\text{LATE} \quad = \quad \frac{\text{2S coef of Q wrt MP}}{\text{1S coef of MP wrt P3}} + \frac{\text{2S coef of Q wrt DV}}{\text{1S coef of DV wrt P3}} \tag{5.12}$$

This LATE was calculated for the individual accounts and tracts, as well as for the group means, based on the WWIV results from models A6 and T6. Prediction error was obtained for the ratio of the estimates by bootstrapping from the estimated coefficient errors, or simulation, over 3,000 iterations. A plot and histogram of the individual LATE calculations by tract appear in Figure 5.15, and these are shown on a map in Figure 5.16. Like the elasticities, the individual LATE calculations show high variability, although the errors are reduced by simulation.

Calculating the LATE from the group means gives a much lower error result, because a higher number of simulations can be used to calculate the error from the ratio of the coefficients. Calculation of the LATE from the account-level models (A4, A5 and A6) uses the 2S coefficients for MP (+1.74, with a variation of 1.05) and DV (11.33 with a variation of 1.87), divides by the 1S coefficients of MP and DV with respect to P3 (23.71, with a variation of 5.86; and -207.85, with a variation of 76.42, respectively), giving a mean LATE of +0.0195 with a sample mean error of +0.0115, using 10,000 simulations for bootstrapping. The *positive* result for the LATE is a result of the positive price elasticity from the second stage model A6, although the contribution of the difference variable significantly reduces the total LATE. Calculation of the LATE from the tract-level models (T4, T5, and T6) uses the 2S coefficients for MP (-2.48, with a variation of 1.64) and DV (1.94 with a variation of 1.95), divides by the 1S coefficients of MP and DV with respect to P3 (22.92, with a variation of 6.74; and -512.31, with a variation of 79.57, respectively), giving a mean LATE of -0.1261 with a sample mean error of +0.001, again using 10,000 simulations for bootstrapping. This is a much more plausible result in sign and magnitude.

5.6 Discussion

The fit of the models grouped by account and tract each have various strengths and weaknesses. The greater number of observations at the account-level improves the group-level model estimates, but the variation around these group-level means results in wildly varying estimates for elasticities and LATE at the individual account level. Tracts remain more efficient at estimating group means, simply because there is more data (and accounts) within each tract, but this does not explain the underlying source of variation within each tract.

One problem in assessing more data and groupings is the computational speed of the analysis. Currently, using existing software, analysis runs for approximately 5% of the data in Chapter 4 takes a week, and larger models have not been resolved. Possibilities for speeding up the calculation include writing a custom-built Gibbs sampler, which is easier to optimize but less flexible for other problems. Parallelization of parts of the calculation may also assist in the speed of the calculation.

One possibility for future work is the addition of joint estimation between the first- and second-stage models using the WWIV method (joint estimation is not possible using the Terza IV because IVs are assigned using the nonlinear rate schedule). Currently the data-level and group-level regressions are estimated jointly in *each* stage, but not in both stages simultaneously. Although the model results are not expected to change because 2SLS estimation is consistent, this would allow errors in the estimation of the IVs to affect the estimation of the data-level predictions. This would lend valuable insight into how IV selection and estimation affects predicted values.

5.7 Tables and Figures

Description	HLM	Terza IV		WW IV		
Stage, Response	ObsMP	1S Q	2S Q	1S MP	1S DV	2S Q
Model code	A1	A2	A3	A4	A5	A6
Average temp	2.10	2.59	2.76	11.47	-151.36	3.98
	(0.14)	(0.18)	(0.18)	(1.11)	(18.10)	(0.23)
Average precip	-0.21	-0.72	-0.48	-2.40	56.27	-0.88
	(0.14)	(0.18)	(0.17)	(1.07)	(18.55)	(0.18)
Sewer	-2.55	-2.04	-2.17	-7.03	126.73	-3.61
	(0.29)	(0.35)	(0.34)	(2.02)	(34.87)	(0.37)
Fixed cost	-2.87	-1.68	-1.48	-1.70	53.75	-3.02
	(0.34)	(0.49)	(0.40)	(3.16)	(52.13)	(0.39)
Indicator 1992	-1.51	-2.59	-2.44	-4.60	41.80	-2.69
	(0.13)	(0.21)	(0.18)	(1.18)	(22.80)	(0.19)
indicator 2001	-0.83	-0.21	-0.69	0.20	-2.21	-1.19
	(0.13)	(0.26)	(0.16)	(1.63)	(30.18)	(0.16)
Intercept *	26.81	22.74	23.00	307.71	-311.57	19.72
	(3.91)	(2.91)	(3.92)	(15.80)	(232.60)	(4.57)
Price, block 1 *		2.94		25.82	-127.14	
		(7.41)		(39.05)	(635.41)	
Price, block 2 *		-2.15		21.01	-26.53	
		(6.34)		(19.77)	(186.42)	
Price, block 3 *		-1.04		4.64	33.22	
		(4.51)		(30.70)	(493.32)	
Marginal price *	8.79		-0.24			4.12
	(7.04)		(7.67)			(7.35)
Diff. variable *	-18.15		-3.53			15.36
	(18.67)		(18.67)			(12.01)
Within-group var						
Residual var	11.37	13.72	14.09	86.10	1,442.42	14.09
	(0.09)	(0.10)	(0.11)	(0.63)	(10.27)	(0.11)
R2 in group						
R2 overall						
Deviance	-15,843	-11,905	-11,345	26,626	85,755	-11,342
	(60)	(59)	(54)	(51)	(44)	(56)

Table 5.1: Data-Level Parameter Estimates for HLM Models Grouped by Account. Numbers and parentheses below represent the mean and standard deviation of simulated values from the posterior distribution, similar to standard errors. All data is center-normalized and all reported values are multiplied times 100 for ease of interpretation. Asterisks (*) indicate that random effects are centered on these fixed effects. Plots and histograms for the random effects appear in Figures 5.1-5.3. Group-level coefficients appear in Tables 5.2-5.7.

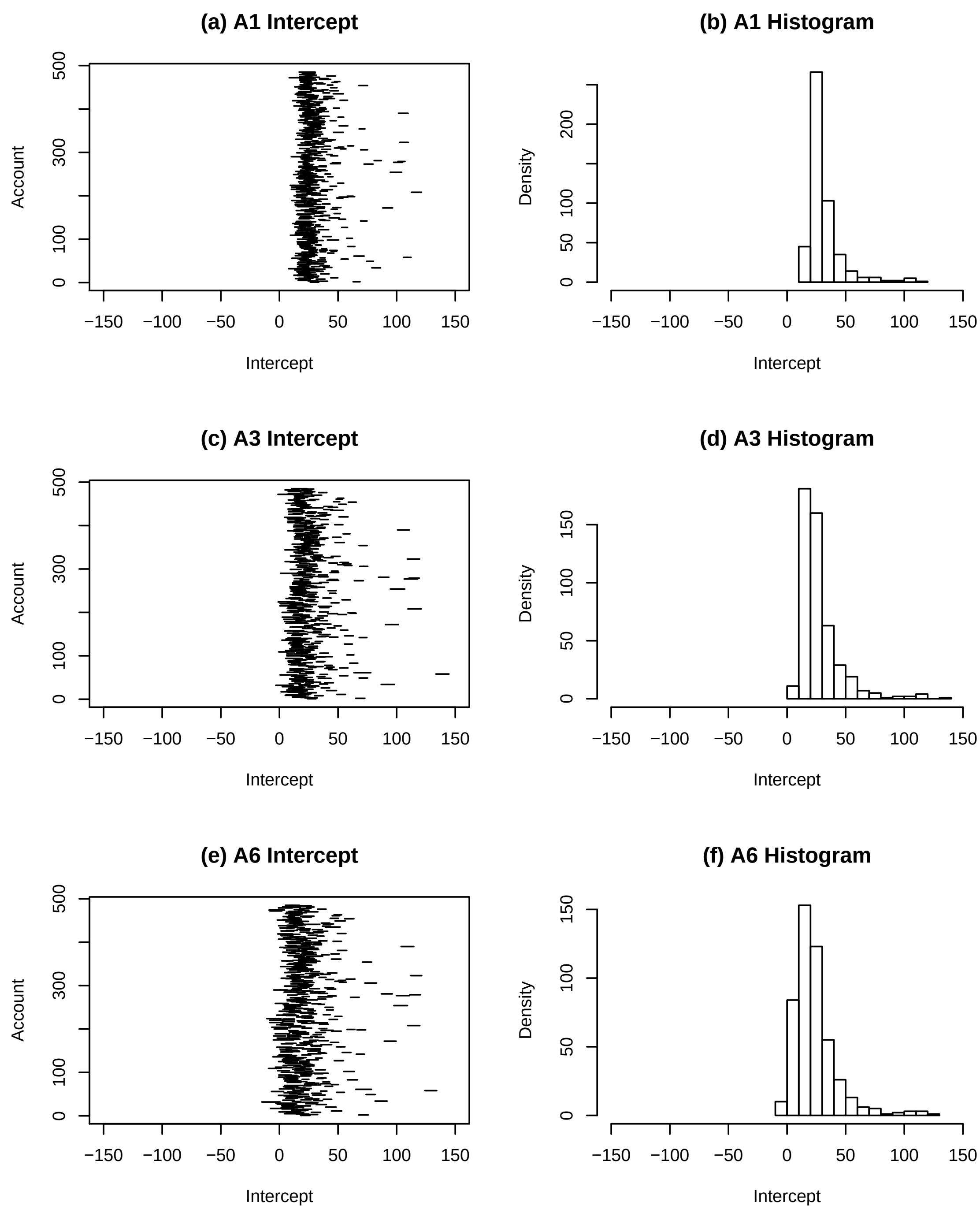

Figure 5.1: Estimated Random Intercepts by Account. All data is center-normalized and all coefficients are multiplied times 100 for ease of interpretation.

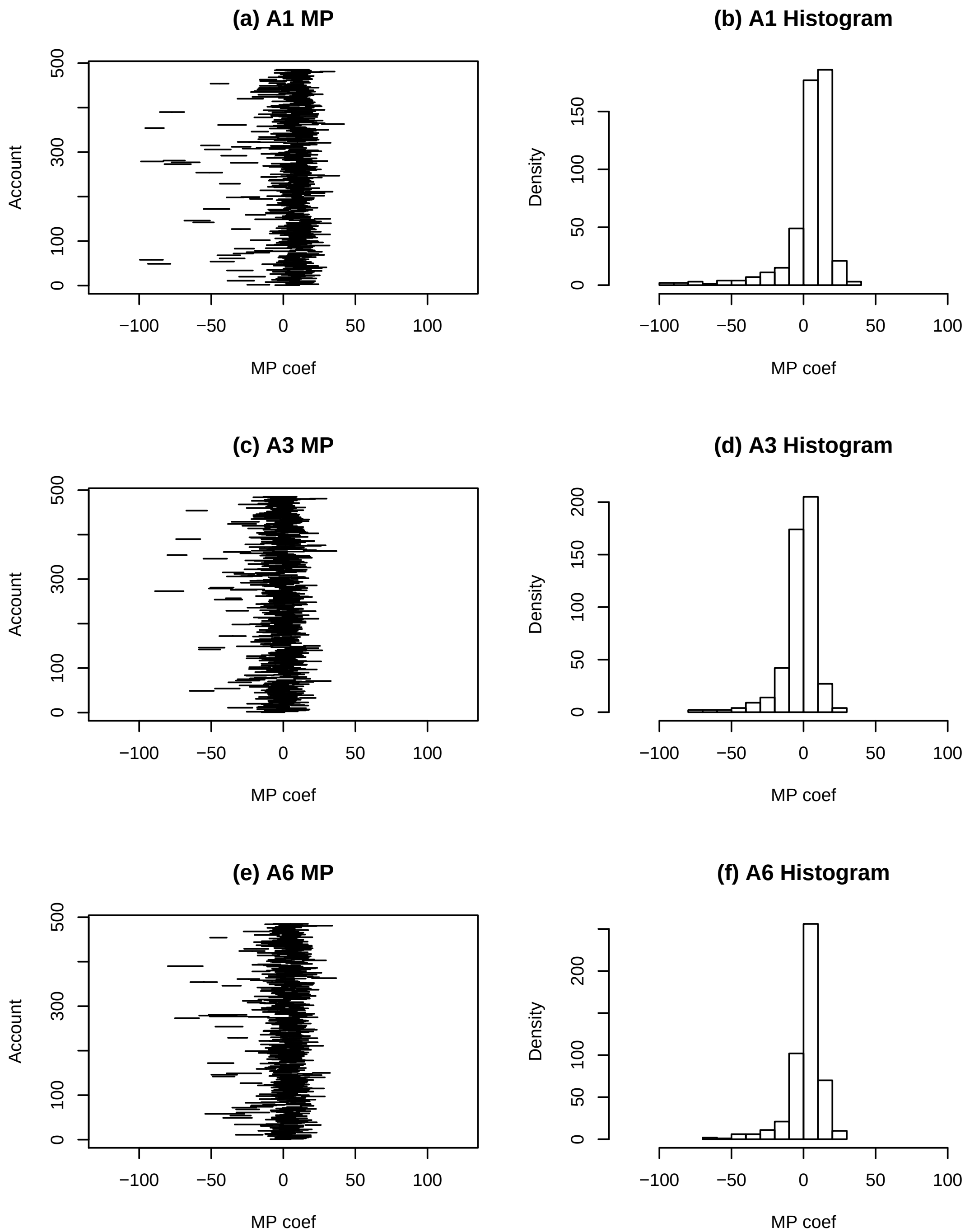

Figure 5.2: Estimated Marginal Price Coefficient by Account. All data is center-normalized and all coefficients are multiplied times 100 for ease of interpretation.

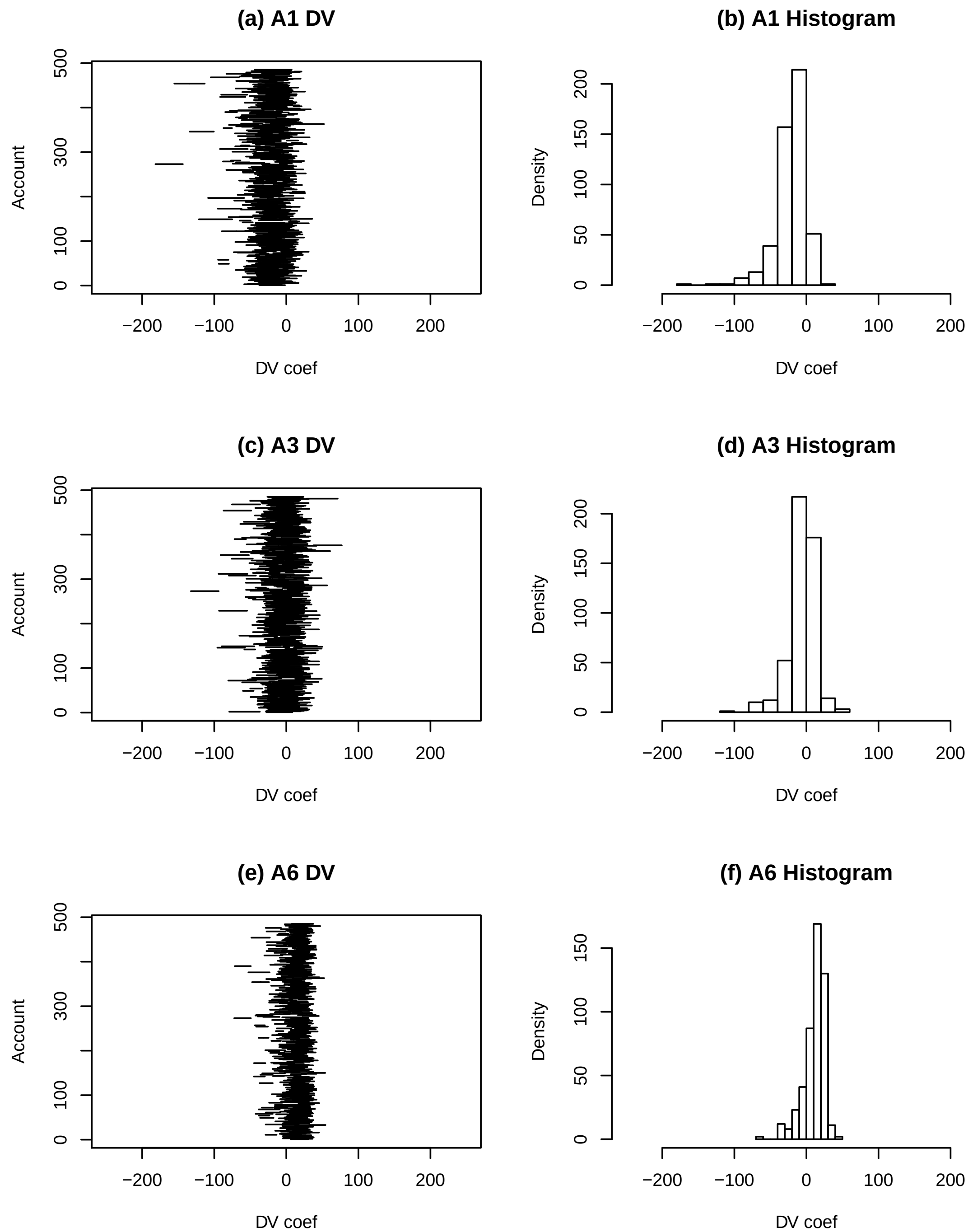

Figure 5.3: Estimated Difference Variable Coefficient by Account. All data is center-normalized and all coefficients are multiplied times 100 for ease of interpretation.

		Group-level γ			
		Intercept	Lotsf	Value	Age
Data-level β	Intercept	31.06	4.17	6.16	-0.14
		(0.71)	(0.71)	(0.67)	(0.72)
	MP	4.43	-5.19	-6.17	0.24
		(1.09)	(0.94)	(0.91)	(0.93)
	DV	-21.43	-3.19	-1.70	0.44
		(2.35)	(1.80)	(1.64)	(1.83)

Table 5.2: Group-Level Estimates by Account for HLM Model A1. Numbers and parentheses below represent the mean and standard deviation of simulated values from the posterior distribution, similar to standard errors. All data is center-normalized and all reported values are multiplied times 100 for ease of interpretation.

		Group-level γ			
		Intercept	Lotsf	Value	Age
Data-level β	Intercept	26.56	4.78	8.05	-0.30
		(0.69)	(0.75)	(0.71)	(0.72)
	P1	1.35	-1.09	-1.66	0.33
		(1.12)	(0.84)	(0.76)	(0.83)
	P2	-1.37	0.74	0.26	-0.05
		(0.75)	(0.80)	(0.66)	(0.74)
	P3	-1.24	-0.80	0.09	-0.10
		(0.66)	(0.48)	(0.46)	(0.47)

Table 5.3: Group-Level Estimates by Account for HLM model A2

		Group-level γ			
		Intercept	Lotsf	Value	Age
Data-level β	Intercept	27.37	5.24	7.79	0.06
		(0.77)	(0.73)	(0.77)	(0.78)
	MP	-2.62	-2.26	-3.88	0.35
		(1.08)	(0.83)	(0.85)	(1.24)
	DV	-5.71	-1.26	-2.82	0.90
		(2.62)	(1.72)	(1.75)	(3.40)

Table 5.4: Group-Level Estimates by Account for HLM model A3

		Group-level γ			
		Intercept	Lotsf	Value	Age
Data-level β	Intercept	320.57	16.12	30.49	1.70
		(2.82)	(2.83)	(2.68)	(2.86)
	P1	13.34	-9.10	-16.12	3.54
		(7.76)	(4.77)	(4.70)	(5.08)
	P2	23.47	-0.65	5.09	-0.70
		(2.74)	(3.23)	(2.73)	(2.99)
	P3	23.71	20.72	27.37	-0.74
		(5.86)	(4.39)	(4.13)	(4.64)

Table 5.5: Group-Level Estimates by Account for HLM model A4

		Group-level γ			
		Intercept	Lotsf	Value	Age
Data-level β	Intercept	-590.76	-236.85	-484.87	-20.21
		(40.23)	(44.72)	(41.73)	(47.25)
	P1	-108.27	24.07	233.94	-91.74
		(84.70)	(60.25)	(54.67)	(43.88)
	P2	-19.25	53.92	-74.45	7.17
		(47.76)	(37.23)	(28.68)	(37.85)
	P3	-207.85	-260.09	-480.79	63.10
		(76.42)	(71.52)	(53.51)	(46.25)

Table 5.6: Group-Level Estimates by Account for HLM model A5

		Group-level γ			
		Intercept	Lotsf	Value	Age
Data-level β	Intercept	23.27	5.64	8.92	-0.42
		(0.86)	(0.83)	(0.83)	(0.85)
	MP	1.74	-2.30	-3.86	0.37
		(1.05)	(0.87)	(0.84)	(0.83)
	DV	11.33	-3.93	-6.17	0.93
		(1.87)	(1.25)	(1.18)	(1.26)

Table 5.7: Group-Level Estimates by Account for HLM model A6

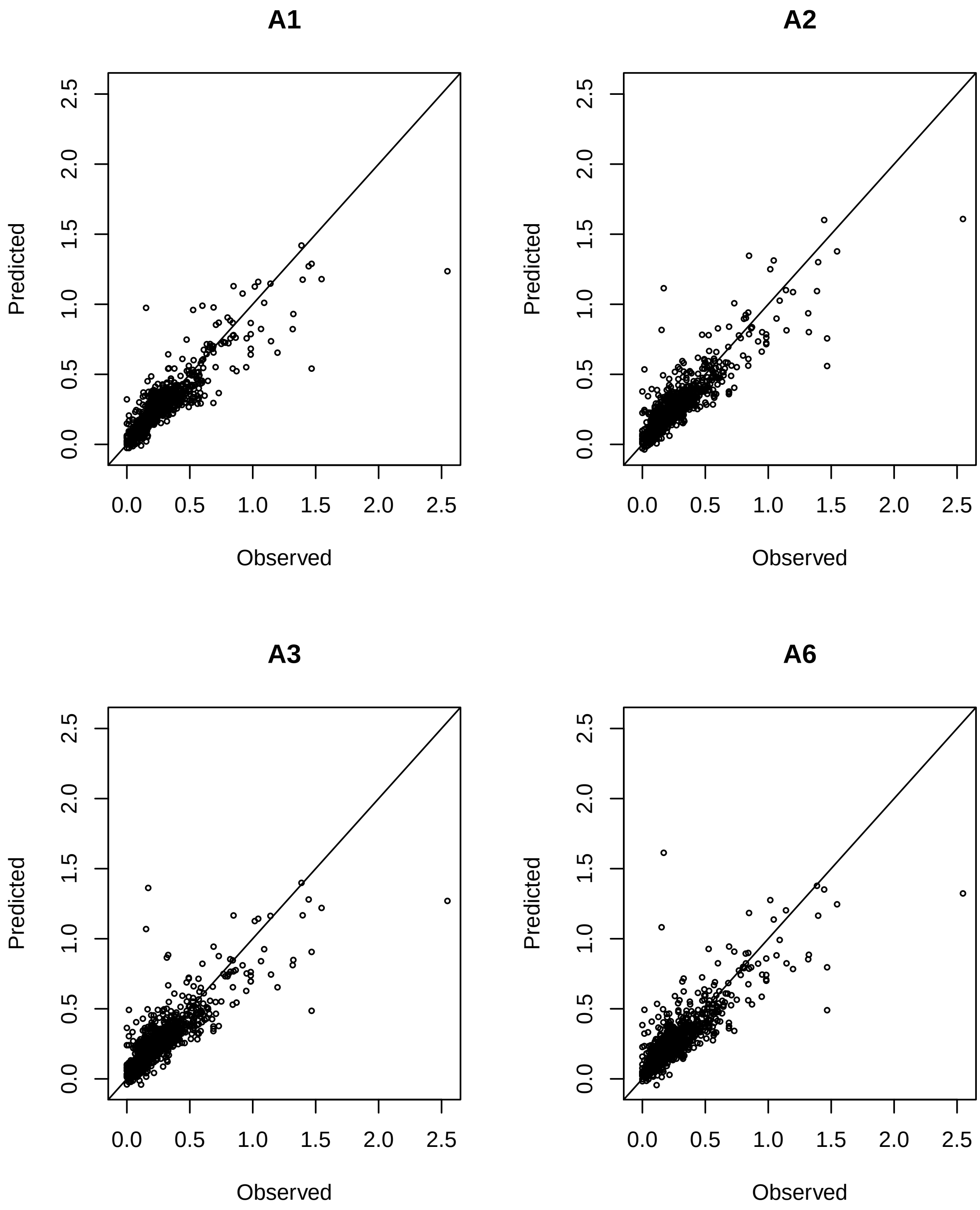

Figure 5.4: Actual versus predicted values for account-level HLM models

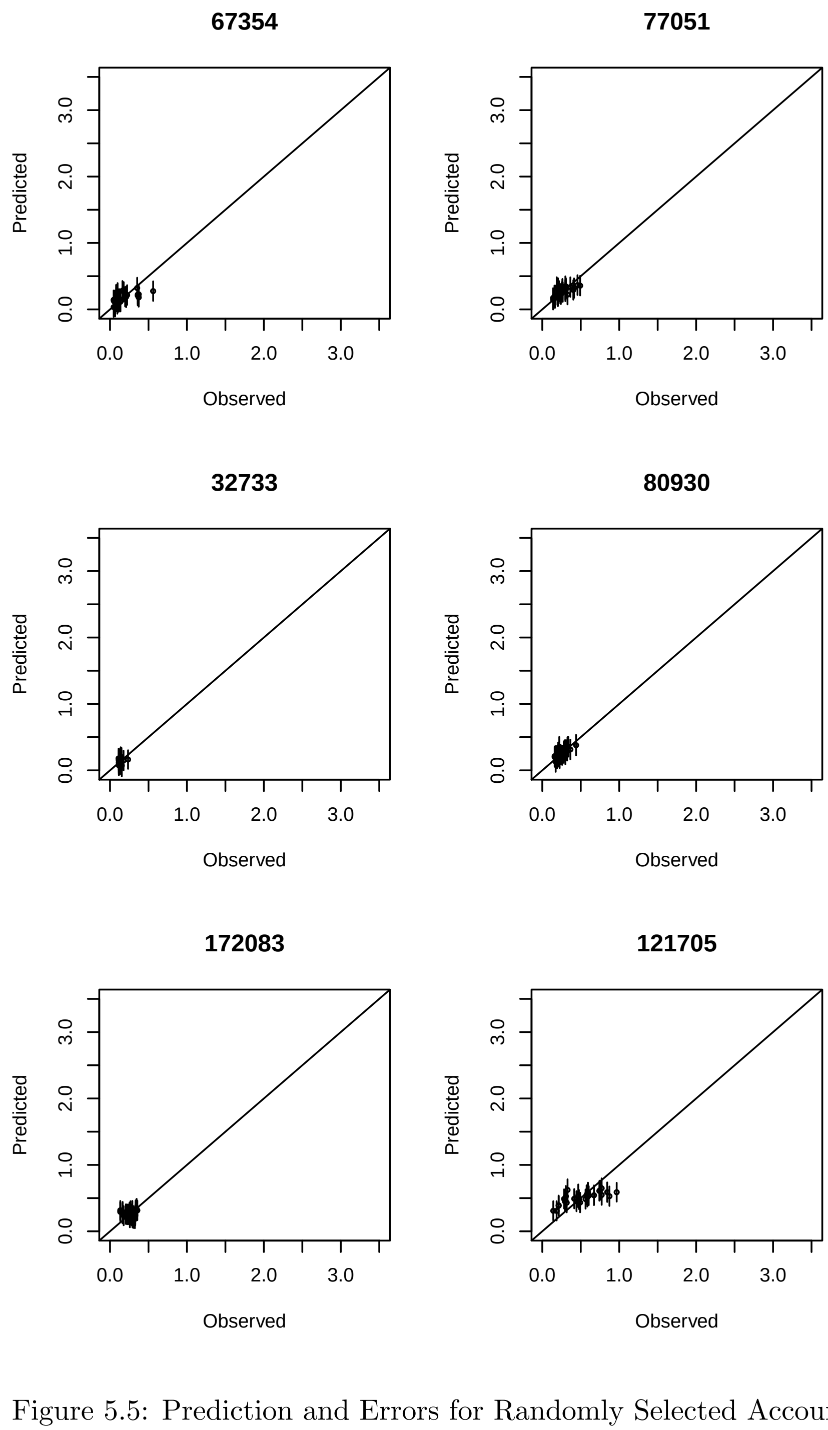

Figure 5.5: Prediction and Errors for Randomly Selected Accounts

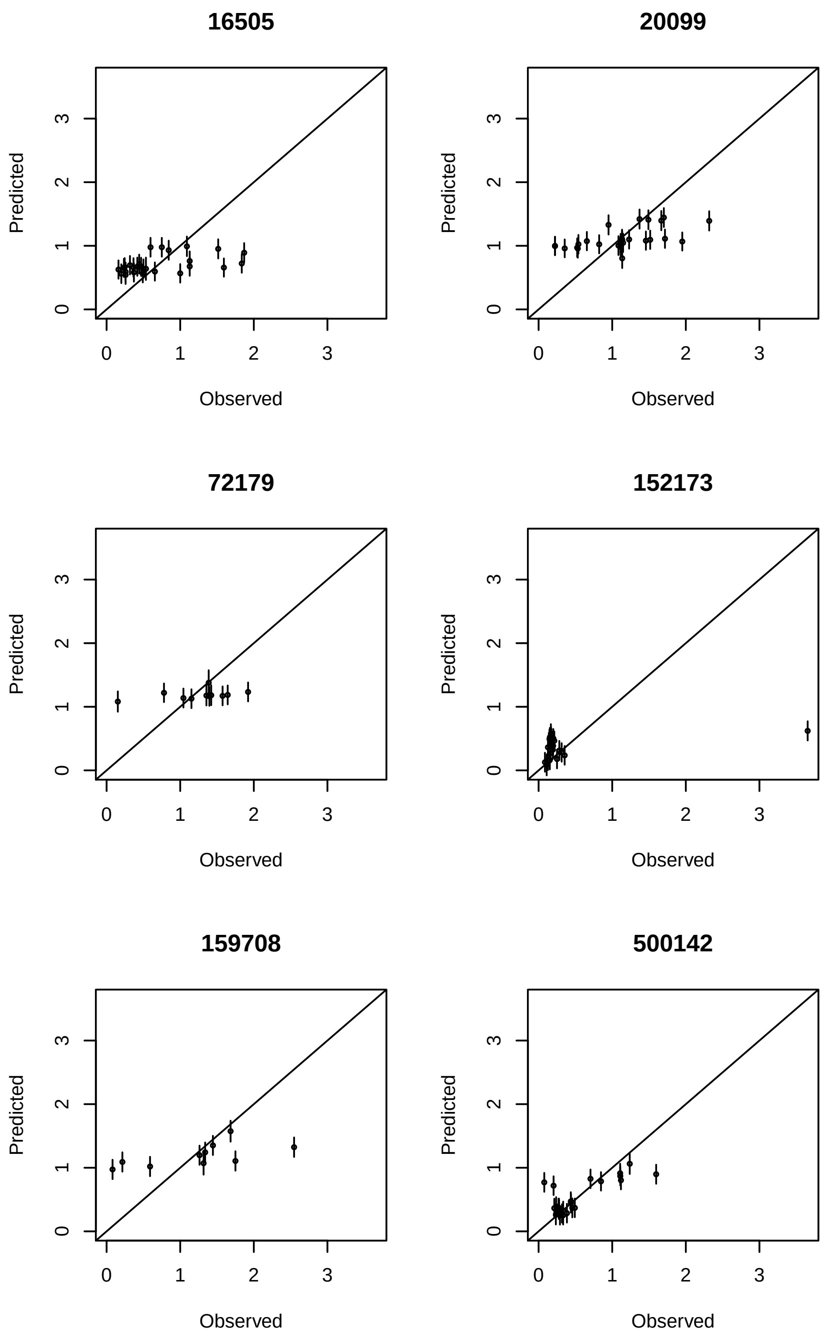

Figure 5.6: Prediction and Errors for Selected High-Consuming Accounts

Description	HLM	Terza IV		WW IV		
Stage, Response	ObsMP	1S Q	2S Q	1S MP	1S DV	2S Q
Model code	T1	T2	T3	T4	T5	T6
Average temp	1.91	2.53	2.96	11.38	-171.31	3.44
	(0.20)	(0.26)	(0.25)	(1.47)	(23.09)	(0.30)
Average precip	-0.04	-0.71	-0.60	-2.95	43.99	-0.79
	(0.20)	(0.26)	(0.24)	(1.44)	(22.50)	(0.27)
Sewer	-3.16	-1.60	-2.10	-5.53	103.46	-2.35
	(0.41)	(0.56)	(0.52)	(2.75)	(44.20)	(0.49)
Fixed cost	-6.56	-1.18	-1.37	-1.53	-52.33	-1.14
	(0.42)	(0.77)	(0.63)	(4.19)	(64.71)	(0.57)
Indicator 1992	-0.77	-2.97	-3.01	-4.64	39.99	-2.96
	(0.18)	(0.31)	(0.29)	(1.60)	(24.24)	(0.27)
indicator 2001	-2.05	-0.27	-0.60	0.56	58.50	-0.78
	(0.18)	(0.40)	(0.22)	(2.24)	(30.95)	(0.23)
Intercept *	26.95	25.05	24.34	312.23	-447.35	24.26
	(1.83)	(2.01)	(2.32)	(10.08)	(141.68)	(2.88)
Price, block 1 *		1.76		25.07	265.44	
		(5.67)		(24.37)	(343.69)	
Price, block 2 *		-1.68		22.93	-121.53	
		(4.71)		(17.79)	(148.73)	
Price, block 3 *		-0.92		4.18	-191.08	
		(3.71)		(18.37)	(264.94)	
Marginal price *	22.33		-2.11			-1.31
	(4.44)		(4.55)			(3.95)
Diff. variable *	5.24		3.72			2.15
	(5.05)		(12.11)			(6.53)
Within-group var						
Residual var	16.32	20.27	20.33	110.92	1,808.58	20.27
	(0.12)	(0.14)	(0.14)	(0.81)	(12.25)	(0.14)
R2 in group						
R2 overall						
Deviance	-8,271	-3,722	-3,652	31,946	90,501	-3,717
	(25)	(27)	(22)	(25)	(22)	(23)

Table 5.8: Data-Level Parameter Estimates for HLM Models Grouped by Tract. Numbers and parentheses below represent the mean and standard deviation of simulated values from the posterior distribution, similar to standard errors. All data is center-normalized and all reported values are multiplied times 100 for ease of interpretation. Asterisks (*) indicate that random effects are centered on these fixed effects. Plots and histograms for the random effects appear in Figures 5.7-5.9. The group-level coefficients appear in Tables 5.9-5.14.

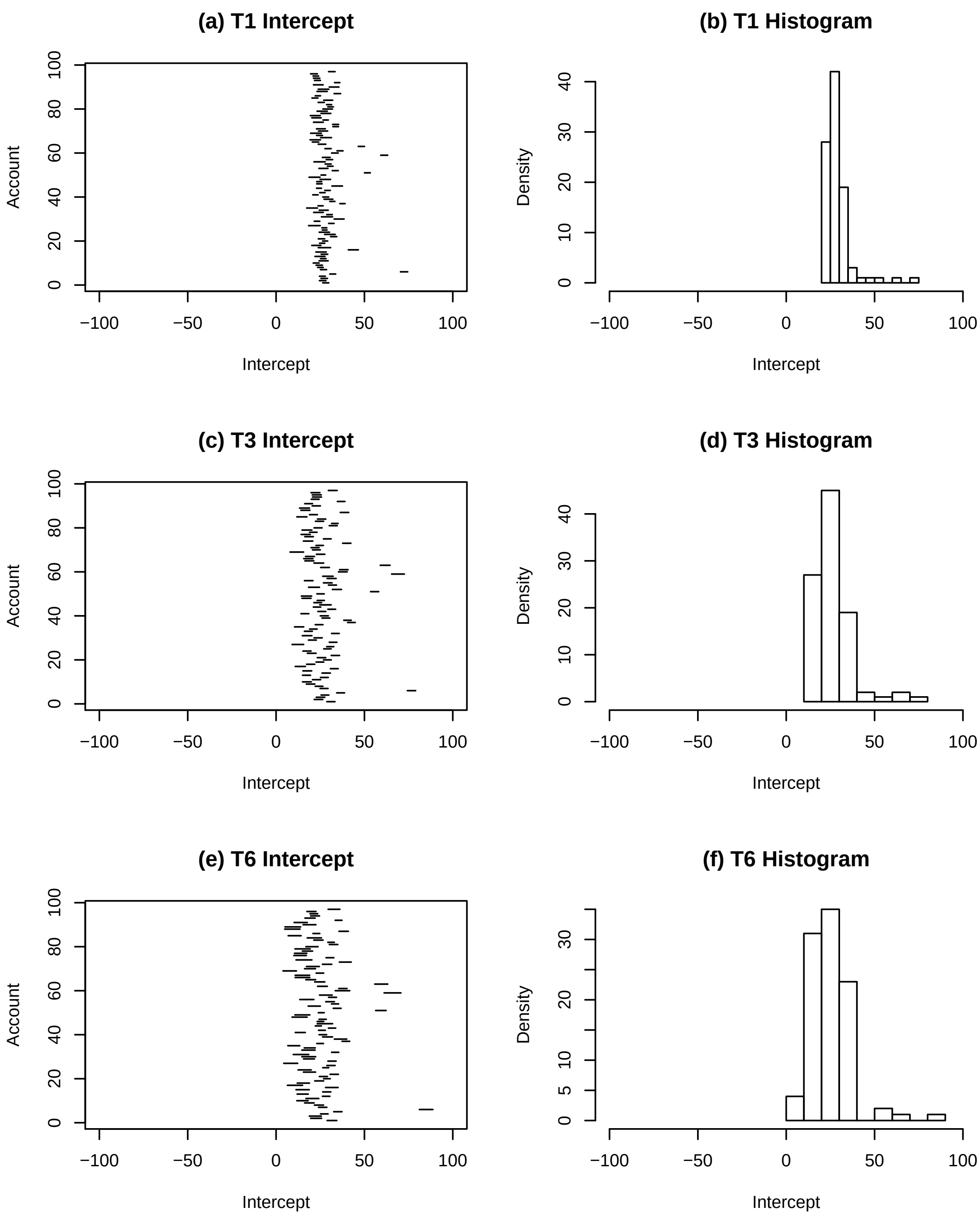

Figure 5.7: Estimated Random Intercepts by Tract.

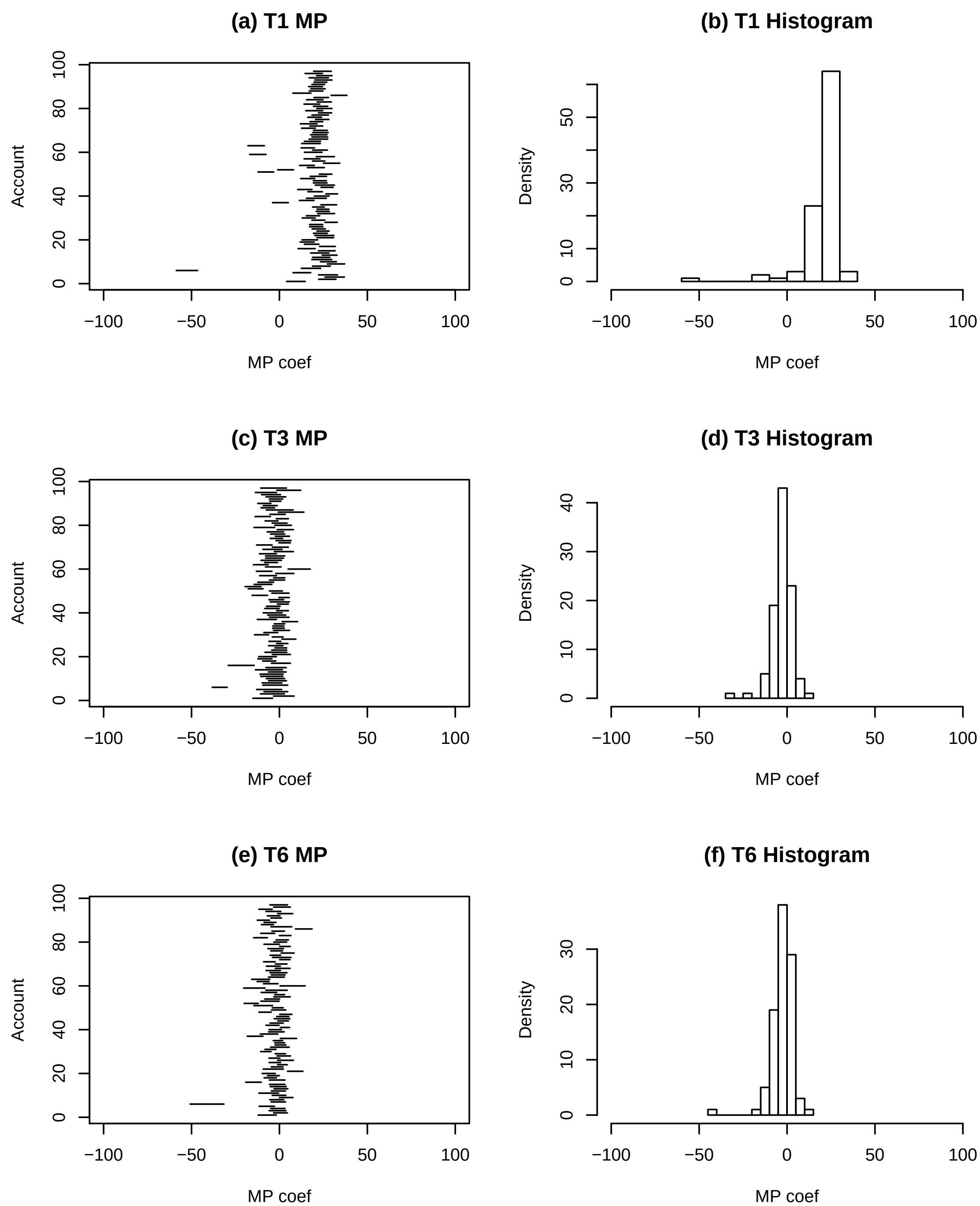

Figure 5.8: Estimated Marginal Price Coefficient by Tract

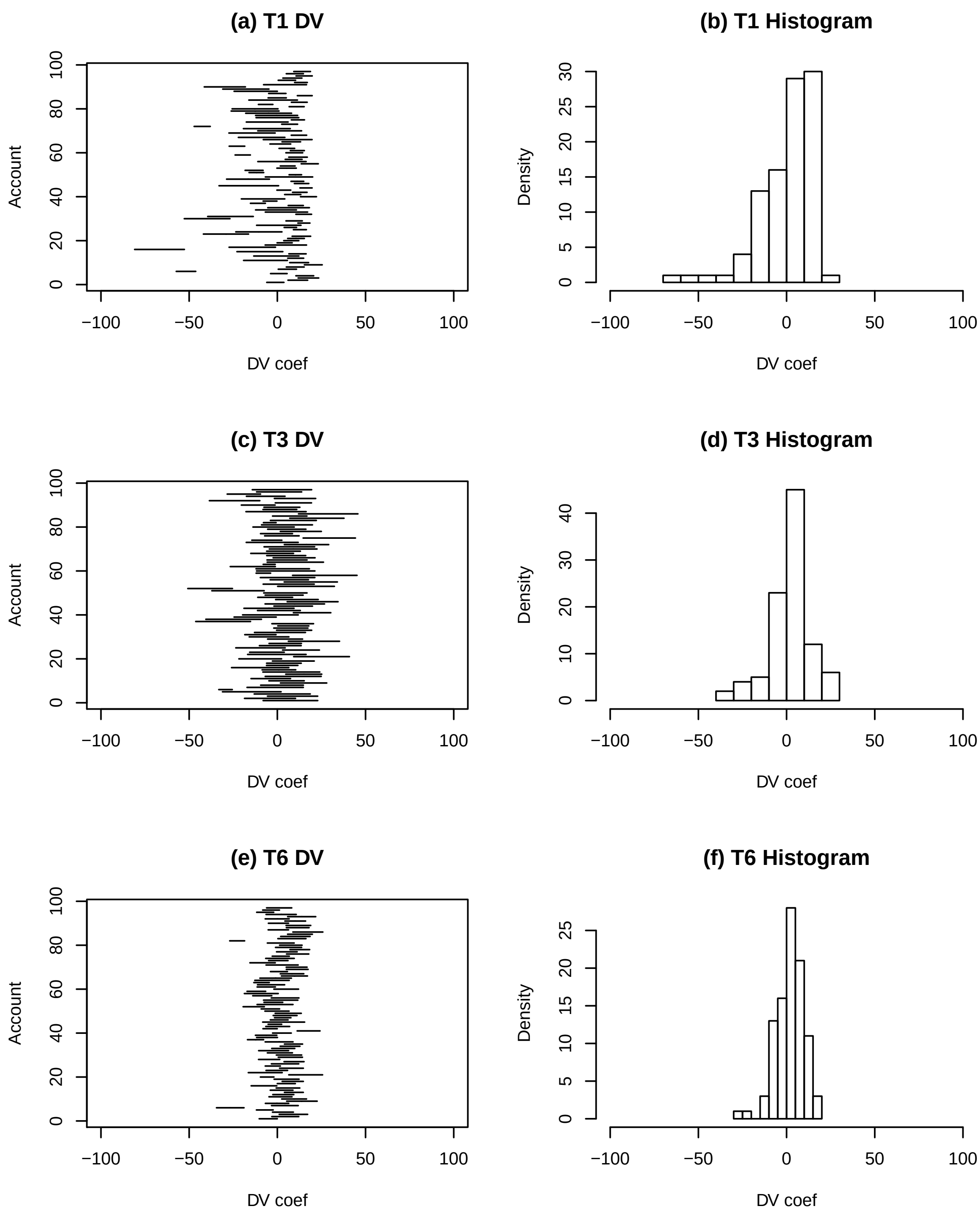

Figure 5.9: Estimated Difference Variable Coefficient by Tract

Data-level β	Group-level γ					
	Intercept	Med income	Av HH size	Med rooms	Med value	Pop
Intercept	28.77	-3.03	-0.72	3.46	4.76	0.67
	(1.28)	(3.70)	(1.82)	(3.05)	(2.38)	(1.38)
MP	20.15	4.68	0.50	-4.29	-6.11	-0.45
	(1.76)	(4.59)	(2.21)	(3.67)	(3.02)	(1.72)
DV	0.17	4.91	0.19	-5.55	-4.20	-2.21
	(2.53)	(6.91)	(3.14)	(5.71)	(4.26)	(2.46)

Table 5.9: Group-Level Estimates by Tract for HLM Model T1. Numbers and parentheses below represent the mean and standard deviation of simulated values from the posterior distribution, similar to standard errors. All data is center-normalized and all reported values are multiplied times 100 for ease of interpretation.

Data-level β	Group-level γ					
	Intercept	Med income	Av HH size	Med rooms	Med value	Pop
Intercept	27.09	-3.39	-0.74	3.38	6.75	0.49
	(1.50)	(4.47)	(2.01)	(3.71)	(2.72)	(1.57)
P1	0.97	0.96	1.40	-1.76	-0.67	-2.05
	(1.94)	(4.24)	(2.20)	(3.75)	(2.60)	(1.66)
P2	-1.75	-0.51	-0.65	1.09	0.06	1.50
	(1.51)	(4.12)	(2.18)	(3.65)	(2.54)	(1.63)
P3	-0.87	2.58	0.13	-2.60	-0.85	0.17
	(1.45)	(3.58)	(1.74)	(2.95)	(2.29)	(1.34)

Table 5.10: Group-Level Estimates by Tract for HLM Model T2

Data-level β	Group-level γ					
	Intercept	Med income	Av HH size	Med rooms	Med value	Pop
Intercept	26.51	-3.68	-1.14	4.22	6.64	0.87
	(1.53)	(4.44)	(2.04)	(3.71)	(2.75)	(1.55)
MP	-3.07	3.54	0.97	-3.24	-1.74	-0.11
	(1.60)	(4.22)	(2.05)	(3.58)	(2.57)	(1.53)
DV	2.43	8.42	3.20	-11.36	-4.31	-3.98
	(3.55)	(8.51)	(3.72)	(6.79)	(5.00)	(2.63)

Table 5.11: Group-Level Estimates by Tract for HLM Model T3

Data-level β	Group-level γ					
	Intercept	Med income	Av HH size	Med rooms	Med value	Pop
Intercept	322.44	-5.49	1.19	2.99	24.91	-0.41
	(3.96)	(10.95)	(5.70)	(9.75)	(7.07)	(4.26)
P1	15.45	7.92	4.23	-0.81	-10.79	0.22
	(8.94)	(14.20)	(7.14)	(12.35)	(8.61)	(5.80)
P2	22.27	4.26	-3.26	-4.10	-2.37	1.90
	(4.51)	(11.73)	(5.39)	(9.85)	(6.98)	(4.15)
P3	22.92	-17.09	1.49	6.88	29.98	-3.52
	(6.74)	(12.92)	(6.84)	(11.64)	(8.12)	(5.38)

Table 5.12: Group-Level Estimates by Tract for HLM Model T4

Data-level β	Group-level γ					
	Intercept	Med income	Av HH size	Med rooms	Med value	Pop
Intercept	-638.76	-77.51	-37.77	102.46	-307.04	22.44
	(56.19)	(97.77)	(70.07)	(86.57)	(60.21)	(60.98)
P1	342.58	0.58	-156.34	-8.89	-59.20	-30.83
	(118.10)	(172.37)	(81.50)	(145.19)	(114.29)	(71.46)
P2	-116.34	-101.09	123.78	29.69	138.91	0.05
	(61.09)	(164.73)	(46.46)	(120.50)	(101.43)	(46.28)
P3	-512.31	100.48	-1.34	15.14	-311.18	65.23
	(79.57)	(113.74)	(69.20)	(102.80)	(98.79)	(67.00)

Table 5.13: Group-Level Estimates by Tract for HLM Model T5

Data-level β	Group-level γ					
	Intercept	Med income	Av HH size	Med rooms	Med value	Pop
Intercept	25.19	-4.18	-2.17	5.14	6.93	0.79
	(1.57)	(4.56)	(2.28)	(3.89)	(2.95)	(1.70)
MP	-2.48	2.33	0.88	-2.75	-2.10	-0.68
	(1.64)	(4.08)	(1.97)	(3.47)	(2.48)	(1.46)
DV	1.94	-0.63	3.29	-2.24	-1.43	-1.43
	(1.95)	(4.94)	(2.27)	(4.05)	(2.94)	(1.77)

Table 5.14: Group-Level Estimates by Tract for HLM Model T6

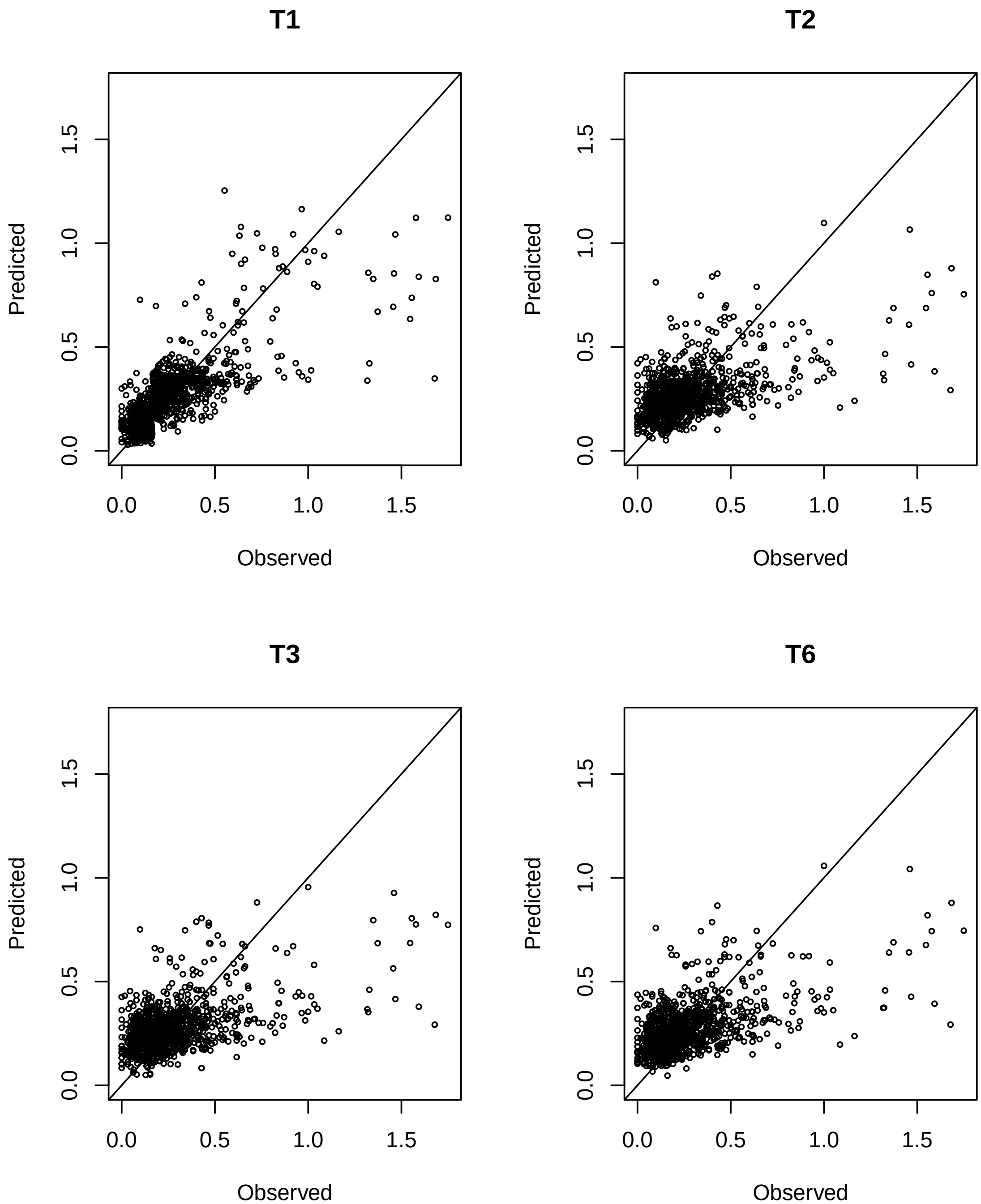

Figure 5.10: Actual versus predicted values for account-level HLM models

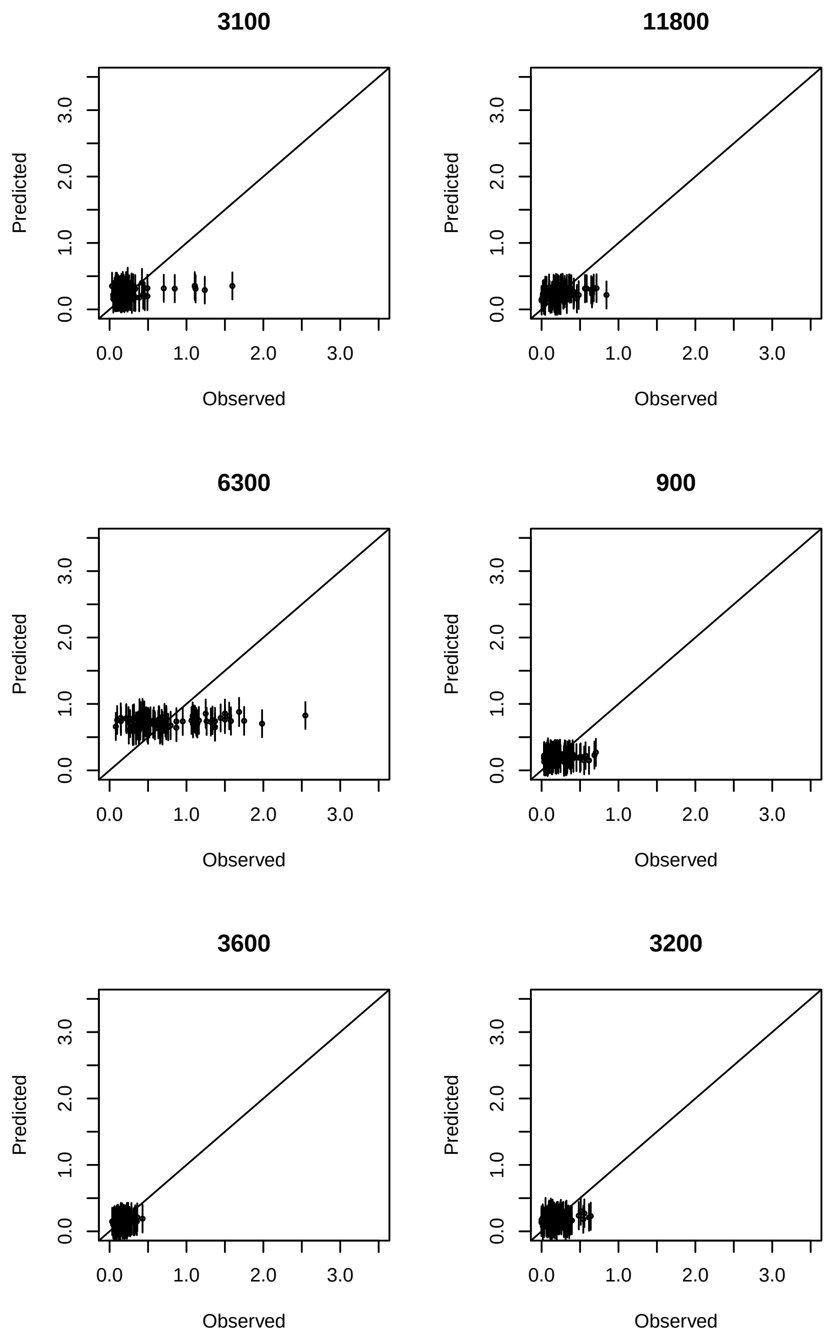

Figure 5.11: Prediction and Errors for Randomly Selected Accounts

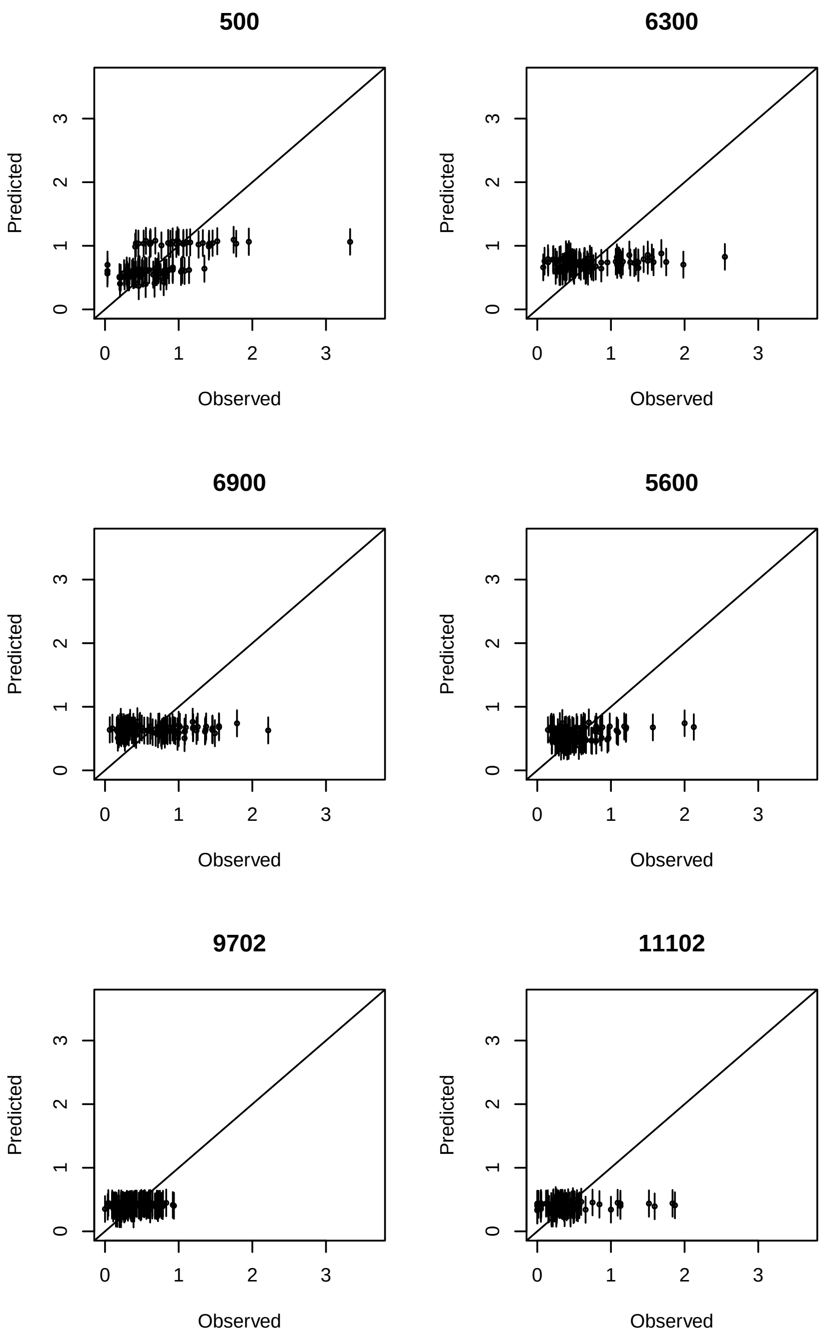

Figure 5.12: Prediction and Errors for Selected High-Consuming Accounts

	Mean P ($)	Mean Q (ccf)	Ratio
Across all tiers	3.15	0.27	11.63
Tier 1	2.48	0.11	22.58
Tier 2	3.15	0.32	9.79
Tier 3	9.75	0.80	12.16

Table 5.15: Table of Average Price and Quantity by Tier. Used to calculate the elasticity by tier below.

Estimates		Calculated elasticities			
Model	MP coef	All tiers	Tier 1	Tier 2	Tier 3
A1	4.43	0.69	1.33	0.58	0.72
	(1.09)	(0.17)	(0.33)	(0.14)	(0.18)
A3	-2.62	-0.41	-0.79	-0.34	-0.42
	(1.08)	(0.17)	(0.32)	(0.14)	(0.17)
A6	1.74	0.27	0.52	0.23	0.28
	(1.05)	(0.16)	(0.32)	(0.14)	(0.17)
T1	20.15	3.12	6.05	2.62	3.26
	(1.76)	(0.27)	(0.53)	(0.23)	(0.28)
T3	-3.07	-0.47	-0.92	-0.40	-0.50
	(1.60)	(0.25)	(0.48)	(0.21)	(0.26)
T6	-2.48	-0.38	-0.74	-0.32	-0.40
	(1.64)	(0.25)	(0.49)	(0.21)	(0.27)

Table 5.16: Estimated Parameters and Calculated Elasticities for an Average Household

	Elasticity predictors			Calculated elasticities				
				Mean	Lotsf		Value	
Model	Mean	Lotsf	Value	HH	+1 sd	-1 sd	+1 sd	-1 sd
A1	4.43	-5.19	-6.17	0.69	-0.12	1.49	-0.27	1.64
	(1.09)	(0.94)	0.91	(0.17)	(0.31)	(0.02)	(0.03)	(0.31)
A3	-2.62	-2.26	-3.88	-0.41	-0.75	-0.06	-1.01	0.19
	(1.08)	(0.83)	(0.85)	(0.17)	(0.30)	(0.04)	(0.30)	(0.04)
A6	1.74	-2.30	-3.86	0.27	-0.09	0.62	-0.33	0.87
	(1.05)	(0.87)	(0.84)	(0.16)	(0.30)	(0.03)	(0.29)	(0.03)

Table 5.17: Sensitivity Analysis for Calculated Elasticities for an Average Household

	Elasticity predictors			Calculated elasticities				
				Mean	Med Inc		Med Val	
Model	Mean	Med Inc	Med Val	HH	+1 sd	-1 sd	+1 sd	-1 sd
T1	20.15	4.68	-6.11	3.12	3.84	2.39	2.17	4.06
	(1.76)	(4.59)	(3.02)	(0.27)	(0.98)	0.44	(0.74)	0.19
T3	-3.07	3.54	-1.74	-0.47	0.07	-1.02	-0.74	-0.21
	(1.60)	(4.22)	(2.57)	(0.25)	(0.90)	0.41	(0.65)	0.15
T6	-2.48	2.33	-2.10	-0.38	-0.02	-0.74	-0.71	-0.06
	(1.64)	(4.08)	(2.48)	(0.25)	(0.88)	0.38	(0.64)	0.13

Table 5.18: Sensitivity Analysis for Calculated Elasticities for an Average Census Tract

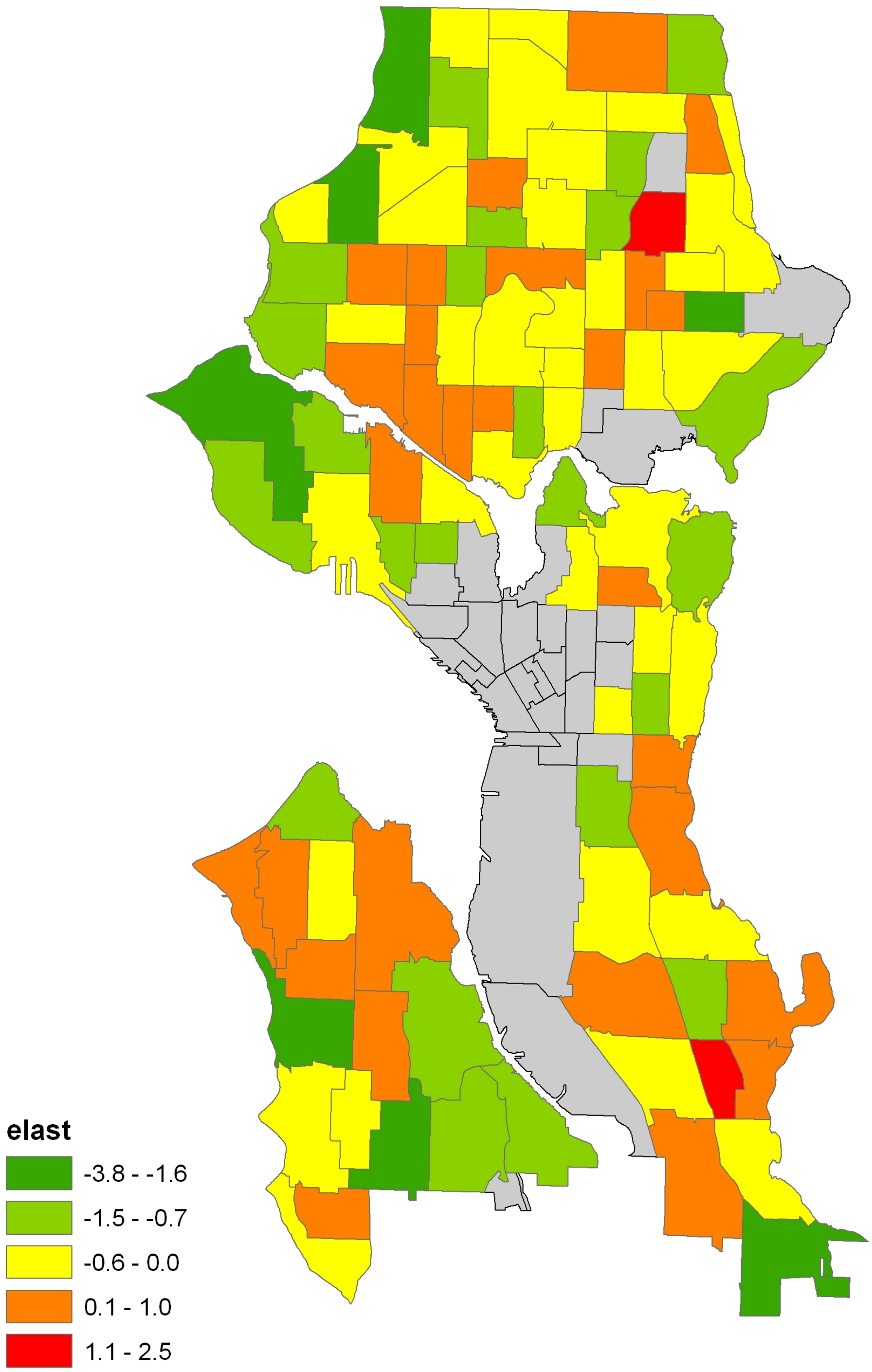

Figure 5.13: Predicted tract-level elasticities from HLM models

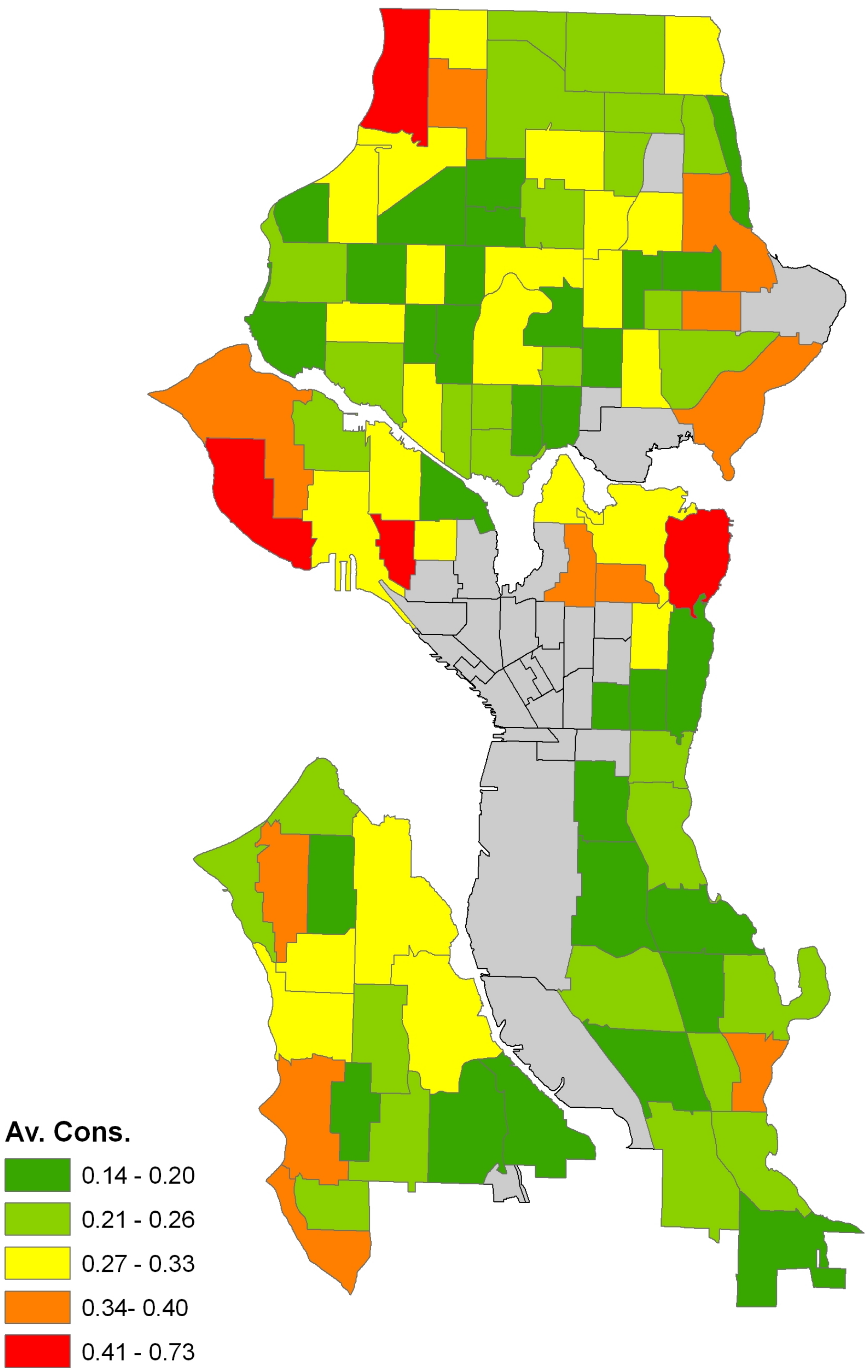

Figure 5.14: Average consumption for census tracts

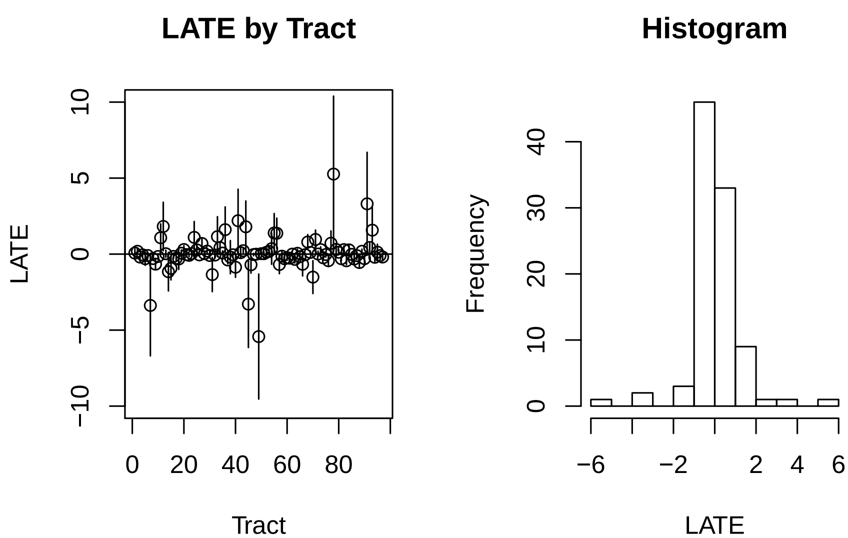

Figure 5.15: Prediction and Errors for LATE by Tract from HLM Models

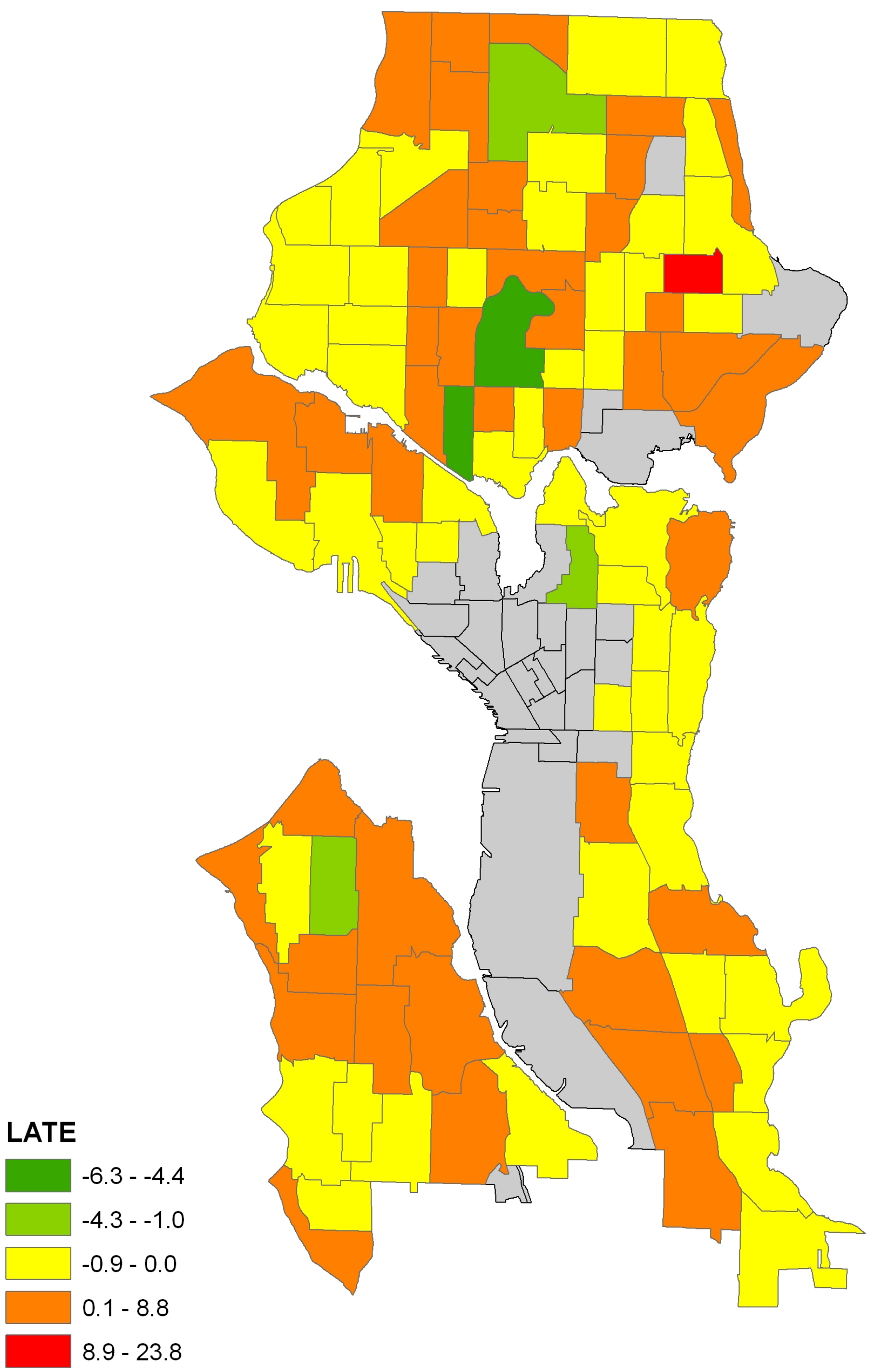

Figure 5.16: Predicted tract-level LATE from HLM models

Chapter 6

POLICY ANALYSIS AND EVALUATION

6.1 Introduction

Chapters 4 and 5 introduced two new statistical methods to modeling of water demand – the CRC and HLM models, respectively – in order to describe better the effects of pricing changes on heterogeneous customers. In order to make these predictive models relevant to policy, we now would like to connect our results to the welfare analysis of proposed policy changes.

This chapter will therefore use the results from the previous chapters to obtain estimates of the impact to welfare of the pricing policy change. I believe that this is a novel contribution to the literature on applied welfare measures, because it opens a number of possibilities for fine-grained policy analysis and assessment. The advantage introduced by the use of these models – in particular, the HLM models – is that it is possible to use the parameter estimates to predict welfare changes and their associated statistical properties at multiple levels. Welfare changes can be calculated at the level of individual bills, households, or groups, and compared to the observed outcomes. It would also be possible to obtain sensitivity analyses of welfare changes based on group-level characteristics. Furthermore, the use of HLM models allow the prediction of water consumption outcomes in alternative situations based on group-level predictors, for example, in a new planned neighborhood that might share similar housing types with other, existing neighborhoods.

A key consideration in the estimation of welfare changes, however, is the propagation of statistical errors from the estimated parameters. Section 2.3 examined previous efforts to measure the statistical properties of welfare measures, when the welfare measures themselves are based on estimated parameters from demand equations.

This chapter will begin by examining previous calculations of welfare from first an economic, and then a statistical perspective, and explore the necessary assumptions to obtain well-behaved distributions for the welfare measures. The common economic and statistical assumptions found will be used to justify the subsequent choice of Bayesian assumptions in the model specification of the demand equation. Finally, using estimated parameters from Chapters 4, 5, and the new model specification, the statistical properties of welfare changes resulting from the pricing policy change will be computed, compared, and discussed.

6.2 Economic Theory

At the outset, we are confronted by several broad questions, as we seek to modeling the welfare changes resulting from the pricing policy change for water. First, do we use a partial equilibrium analysis, in which the effects on only the market for water are concerned; or do we use a general equilibrium analysis, where the market for water is considered in relation to the rest of the broader economy and its interrelations? Second, and related to the first question, is it reasonable to use a single demand equation for water, or does it need to be part of a broader set of linear expenditure functions? Third, what welfare measure should we use? A great deal has been written about all of these issues in the economic literature.

For the more limited purposes of this chapter, we will take the simplest option, using a single demand equation to model water demand in a partial equilibrium analysis. The justification for this is that although water is a critically important resource for the environment and public health, water as an economic good constitutes neither a major portion of household budgets nor the larger economy. Water consumption is a relatively small fraction of most household budgets, and water used for drinking water is a small fraction of the overall amount of surface water withdrawals in the United States (Kenny et al., 2009), and therefore unlikely to affect the broader economy. Finally, the reliability of drinking water systems is rarely an issue in the United States and much of the developed world.[1]

Furthermore, in order to measure the welfare change as a result of the pricing policy change, this section will explore two alternative measures of impact to consumer surplus: the Marshallian consumer surplus, and the exact CS calculation introduced by Hausman (1981).

---

[1] Although concerns are growing over the issue of water quality and contamination from decaying existing infrastructures (Duhigg, 2009a,b).

6.2.1 Marshallian Consumer Surplus

The simplest calculation for welfare changes as a result of the price change is the Marshallian CS, the lost CS to the left of the Marshallian demand curve. Figures 6.1- 6.4 show how the Marshallian CS is calculated. Figure 6.1 shows that various consumers have demand curves varying by intercept and slopes. The consumer and producer surplus are defined in Figure 6.2. When prices are changed exogenously, then distribution of Marshallian CS is described by Figure 6.3, and the loss in Marshallian CS is described by the shaded area in Figure 6.4.

Calculating the Marshallian CS is relatively straightforward, as the area of the trapezoid in Figure 6.4:

$$\triangle CS = \frac{1}{2} \times (p_2 - p_1) \times (x_2 + x_1) \tag{6.1}$$

6.2.2 Exact Consumer Surplus

I will also use the approach introduced by Hausman (1981) to calculate the exact change in consumer's surplus (CS). Previously, the analysis and bounds in Willig (1976) were used to justify CS as an appropriate approximation to welfare measures such as contingent variation (CV) and equivalent variation (EV). Freeman (2003, page 49, 69) gives an overview of the different welfare measures, and surveys previous approaches to approximate values of CS, CV, and EV. Slesnick (1998) reviews alternative empirical measures of welfare, but for all of the reasons discussed above, many of the issues in this article do not arise in water demand, such as multiple price changes; nonlinear terms in the expenditure function; and large income effects. A more complicated derivation for the welfare change is Hausman's calculation of CV, which has the advantage of being an exact calculation that does not require approximation. This will now be presented below.

In the case of water demand, we seek to model the demand for a single good relative to all other goods. We can generalize to this by starting with a consumer's choice between an arbitrary number of goods n. We will also derive the indirect utility functions related to our model specifications, because these functions will be useful in calculating the welfare implications of changes in key variables.

A consumer is faced with the problem of maximizing utility U over n goods at $x = (x_1 \dots x_n)$ quantities subject to a budget constraint,

$$\max_x U(x) \quad \text{subject to} \quad \sum_{i=1}^{n} p_i x_i = px \leq y \tag{6.2}$$

where p is the vector of prices for the n goods and y is the consumer's total income. The indirect utility function $V(p, y)$ is defined as the solution to this maximization problem,

$$V(p, y) \equiv \max[U(x) : px \leq y]$$

Using duality theory, the indirect utility function has an important property governing the rate of change of maximum utility with respect to a price, also known as Roy's identity, which can be derived using the envelope theorem (Silberberg and Suen, 2001):

$$x_j = -\frac{\partial V(p, y)/\partial p_i}{\partial V(p, y)/\partial y}$$

Now, in order to justify the specification and estimation of a consumer demand function with a single observed price, we need to make several further assumptions as suggested by Hausman (1981). First, we only assume that here are only two goods, the quantity of water x_1, and all other goods aggregated together into a group $g(x_2 \dots x_n)$. Second, the initial respective prices for the goods are $p = (p^0, 1)$, where the group of all other goods serves as the numeraire with $p_2 = 1$, and all prices and incomes are therefore deflated with respect to the price of the second good. Third, we assume a separable utility function such as $u(x_1 \dots x_n) = u(x_1, g(x_2 \dots x_n))$.

Using these assumptions together with Roy's identity, we can derive the appropriate utility functions that give us the desired functional forms for our observed market demand curves. For example, the form of the utility function can be obtained from the common linear regression equation in the following manner. Setting the observed consumer demand function equal to Roy's identity,

$$\begin{aligned} x_1 &= \alpha p_1 + \delta y + \gamma z \\ &= -\frac{\partial V(p,y)/\partial p_i}{\partial V(p,y)/\partial y} \end{aligned} \tag{6.3}$$

This linear partial differential equation can be solved by applying the method of characteristic curves. For a given indifference curve with constant utility u_0, as the price changes the utility level remains constant:

$$V(p_1(t), y(t)) = u_0$$

Applying the implicit function theorem along the path of the price change and on the path of constant utility yields no change with respect to prices and income,

$$\frac{\partial V(p_1(t), y(t))}{\partial p_1(t)} \frac{dp_1(t)}{dt} + \frac{\partial V(p_1(t), y(t))}{\partial y(t)} \frac{dy(t)}{dt} = 0$$

Using Roy's identity, we can then solve the following ordinary differential equation

$$\frac{dy(p_1)}{dp_1} = \alpha p_1 + \delta y + \gamma z$$

in order to obtain income y as a function of price p_1

$$y(p_1) = ce^{\delta p_1} - \frac{1}{\delta}\left(\alpha p_1 + \frac{\alpha}{\delta} + z\gamma\right)$$

and therefore solve for the indirect utility function v_{lin} for the linear form:

$$v_{lin}(p_1, y) = e^{-\delta p_1}\left[y + \frac{1}{\delta}\left(\alpha p_1 + \frac{\alpha}{\delta} + z\gamma\right)\right] \tag{6.4}$$

Crucially, Hausman (1981) notes that this utility function is restricted to be decreasing in prices if $\alpha \leq 0$ and increasing in income if $\delta \geq 0$.

The compensating variation (CV) is defined as the amount of money required to keep the customer at the same initial utility level, that is,

$$\text{CV}(p_0, p_1, y_0) \quad = \quad e(p_1, u_0) - e(p_0, u_0) \tag{6.5}$$

where e is the associated expenditure function dual to the utility being maximized in equation 6.2. Combining equation 6.4 with the definition of CV in equations 6.5 allows it to be computed exactly:

$$\text{CV}(p_0, p_1, y_0) \quad = \quad \frac{1}{\delta} e^{\delta(p_1 - p_0)} \left[x_1^0(p_1^0, y_0) + \frac{\alpha}{\delta} \right] - \frac{1}{\delta} \left[x_1^1(p_1^1, y_0) + \frac{\alpha}{\delta} \right] \tag{6.6}$$

We will now examine the statistical theory governing how to compute this quantity with estimated parameters.

6.3 Statistical Theory

In Chapters 4 and 5, we estimated the linear form of the demand equation to obtain parameter values for the price elasticity α, income elasticity δ, and associated errors. Since α and δ are estimated statistically, they are considered to be random variables, and in turn, the CV as a welfare estimator is also a random variable. Upon examination, we can see that a critical quantity determining the value of CV is the ratio of the parameters α/δ. Kling and Sexton (1990) writes, "in general, the mean and variance for a ratio of random variables cannot be expressed analytically in terms of the moments of the two random variables. In fact, the mean and variance for the ratio may not exist". Similarly, Casella et al. (2002) write, "taking ratios can lead to ill-behaved distributions" (p. 108).

Nonetheless, we *can* compute analytically the special case of the ratio of two standard normals. In fact, the ratio of two standard normal random variables X, Y can be computed using the continuous bivariate transformation $U = X/Y$ and $V = |Y|$. The pdf of U can then be shown to be that of a Cauchy random variable:

$$f_U(u) = \frac{1}{\pi(u^2 + 1)}, \qquad \infty \leq u \leq \infty$$

Appendix B shows the full derivation.

This example is particularly important because it demonstrates several inherent difficulties in using a linear functional form with classical standard errors to calculate compensating variation, which is the most likely scenario in applied welfare analysis. First, theoretically the ratio of two normally distributed random variables does not result in a normally distributed random variable. Second, the Cauchy distribution has neither a mean nor a variance, which calls into question the reporting of mean welfare changes based on Hausman's formula. Most importantly, in order to achieve a bivariate transformation, it was necessary to impose the inequality restrictions on X and Y:

$$(X, Y) \in [-\infty \leq x \leq \infty, -\infty \leq y \leq \infty] \rightarrow (U, Y) \in [-\infty \leq u \leq \infty, 0 \leq v \leq \infty]$$

As the full derivation shows, these inequality restrictions are necessary to ensure a continuous, one-to-one transformation from the space of (x, y) to (u, v). Without this restriction, the estimated ratio is not identifiable, because both (x, y) and $(-x, -y)$ both map onto the same point in (u, v) space. These inequality restrictions also mimic one of the key *economic* restrictions that Hausman made on the form of the welfare function equation 6.4, that the coefficient of income δ is greater or equal to zero. Although statistical theory does not place a restriction of the sign of the coefficient of price, $\alpha \leq 0$, this restriction is also compatible with the bivariate transformation of price.

A similar possible restriction that was explored was to incorporate the inequality restrictions directly into the Bayesian estimation process, in the prior selection step for the parameters. A proper prior distribution $P(\theta)$ can be chosen such that the parameters themselves are restricted to the space that is justified by economic and statistical theory; this suggests the use of one-sided distributions that require that $\alpha < 0$ and $\delta > 0$. Furthermore, because of the dependence on CV on $1/\delta$ in equation 6.6, we not only need $\delta > 0$, we need δ values that are not small, to avoid blowing up our welfare measures.

In subsequent analyses, however, the problem found with this method was that with a large number of groups, it is highly likely that there will be a *single* parameter value δ that does *not* obey the one-sided (or truncated) restrictions on the parameter. This then makes it impossible for the Gibbs sampler to simulate conditional probabilities in one of the steps, breaking the iterative chain calculations necessary. For example, out of the 485 accounts in the smaller subsamples, if only *one* account has a positive elasticity, and all of the simulated parameter estimates for that account come back positive, then Gibbs sampling software is unable to simulate any further steps. In practice, parameter restrictions in the estimation process results in very fragile estimation processes.

This section has reviewed the various grounds provided by both economic and statistical theory, to justify restricting the estimated parameters to negative values for the coefficient of price α, and positive values for the coefficient of income δ. The next section will show how these parameter restrictions were applied to the data to calculate the statistical properties of the welfare measures.

6.4 Model Specification & Estimation

We can now turn to estimating the statistical properties of the various welfare measures based on the underlying estimated parameters.

Estimating the errors of the Marshallian CS is relatively straightforward. In equation 6.1, the only source of error are the predicted water consumption levels with and without the price change.

As discussed in section 2.3, with more complicated welfare measures such as Hausman's exact calculation for CS, the literature has suggested several methods to estimate the statistical properties of welfare measures without any conclusion. Adamowicz et al. (1989) were the first to point out that the statistical properties of the welfare measures depend considerably on the functional form chosen in the initial estimation. To estimate the mean and standard deviation, they use a bootstrapping method similar to that of Krinsky and Robb (1986). However, as Kling and Sexton (1990, page 410) note,

> "An apparent problem in the [Adamowicz et al. (1989) analysis] is that the bootstrap-generated data occasionally produced implausible welfare estimators. Specifically, draws generating absolutely small negative values for [δ] yield very large values of willingness to pay, quite possible in excess of income. Positive values for [δ] may also be generated, implying negative willingness to pay. Such

> results are counterintuitive and inconsistent with the underlying economic theory. Yet because these implausible estimators may be very large in absolute value relative to the magnitude of the other bootstrap-generated welfare estimates, they can easily dominate calculations of bias and also cause bootstrap-estimated variances of welfare to be very large."

They therefore employ a modified form of the bootstrapping process, where they apply an improper prior that simply eliminates any welfare estimates that are negative or greater than income, and allows parameter ratios outside of the quadrant allowed by economic theory. The improper prior $P(\theta)$

$$P(\theta) = \begin{cases} 1 & \text{if } 0 < CV_i < Y_i \text{ and } \alpha/\delta < 0 \text{ for } i = 1 \ldots n \\ 0 & \text{otherwise} \end{cases}$$

is simply placed into the posterior density equation

$$P(\theta|X) \propto P(\theta)L(\theta|X)$$

or, that is,

$$\text{posterior} \propto \text{prior} \times \text{likelihood}$$

Figure 6.5 shows the restricted parameter space for α and δ. This is the method that will be used principally in this chapter, because it is relatively robust, and can be applied to existing simulation results.

6.5 Results

The results for the Marshallian CS are relatively straightforward and only require the predicted consumption with and without the imposed price change. The change in Marshallian CS is found to range from approximately 5-45% of the existing average total bills for various census tracts. The tract-level results in Figure 6.6 are based on the water consumption predictions from the model CRC-2T, alongside the predicted tract-level elasticities from Chapter 4. The CRC model was used because it has relatively low prediction errors for water consumption, although as argued in Chapter 4, these errors may not reflect the full errors introduced by getting estimates from statistical distributions with conditional means.

We then tried to calculate the exact CV using the results from the HLM models in Chapter 5. Equation 6.6 gives us an exact calculation for the compensating variation as a function of the change in prices (p_1 and p_0); model predictions (x_1 and x_0); and most importantly (and uncertainly), the estimated parameters for price and income elasticity (α and δ). Since our analysis in Chapter 5 gives us probability distributions for every parameter, we can simulate results for every single one of our observations and groups. In order to summarize the results meaningfully, however, I will aggregate the results in two different ways.

First, the simulation results from Chapter 5 were used to obtain a CV estimate for each observation. The CV estimates were then matched to the original calculated total bills and observed tiers. The simulation results for CV are then aggregated and summarized by tier, as shown in Table 6.1.

Table 6.1 shows that the observed average monthly bills vary quite dramatically between the first two tiers, and the third tier. The first set of calculations are based on simulating the welfare measures and associated errors directly from the parameter estimates, without any parameter restrictions, as suggested by Krinsky and Robb (1986) and Adamowicz et al. (1989). However, while the calculated *mean* CVs without any parameter restrictions are quite reasonable – on the order of 0-50% of the total bills – the variation in the estimates dwarf the mean estimates and are typically much larger than even the observed total bills. For example, in the third tier for simulated welfare measures, the standard deviations range from approximately \$150 - 1150. These large errors associated with the estimated CVs means that the results for the welfare impacts by census tract are inconclusive.

The next set of lines then show the results of the aberrant bootstrapping approach suggested by Kling and Sexton (1990). Only the parameter estimates that result in a welfare change between zero and the total bill are allowed by imposing an improper prior, justified by the fact that willingness to pay (WTP) is bounded by a consumer's total income. While the standard deviation of the welfare measures are significantly reduced to a reasonable order of magnitude, ranging from \$11-12, unfortunately, these errors still outweigh the mean welfare calculations, and are therefore inconclusive.

Second, because of the large variation in the estimated parameters at the data-level (i.e., in the price elasticties by group), I then tried to obtain simulated welfare changes directly from the group-level parameters. That is, for each group, the mean price and income elasticities and their associated variances in equation 5.2 can be calculated directly by using the group-level parameters ($\boldsymbol{\Gamma_j}$) and multiplying by the group-level predictors ($\boldsymbol{u_j}$). The groupings are by household in the account-level models A1, A3 and A6, and by census tract in the tract-level models T1, T3, and T6. Using the average predicted consumption for each group, we can then calculate what the welfare change would be for each household.

The results for the estimated parameter approach are shown in Table 6.2 and Figures 6.7 and 6.8. In this table, the first line shows the observed average monthly bill for a typical group. Similar to previous results and Table 6.1, the welfare measures simulated directly from the welfare measures show extremely poor statistical properties. The mean welfare changes fluctuate by two orders of magnitude, from -\$50 to \$1,575, while the standard errors in each group range wildly from \$5,000 to \$75,000 (!). Again applying the aberrant bootstrap approach proposed by Kling and Sexton (1990), we find that the imposition of parameter restrictions again significantly reduces the estimates for mean welfare and associated errors. Unfortunately, again, the standard deviations for each group typically outweigh the mean welfare changes estimated. A further downside of this second approach is that it is difficult to summarize for many groups the general impact of the price change *when* they fall into the third tier, because many particular groups (households or tracts) never consume water in the third tier.

6.6 Discussion & Future Work

The most obvious problem with the applied welfare measures calculated are the large uncertainties in the estimated parameters from chapter 5. These errors in turn propagate forward into the calculation of applied welfare changes, giving us large errors in the welfare measures and therefore inconclusive results.

As suggested in the conclusion of chapter 5, the main problem in obtaining better estimated parameters is the computational cost. It is possible to obtain much more accurate estimates using approximately 20 times more data, as is demonstrated in chapter 4. Although we do possess the additional data needed, it is difficult to fit the substantially more complicated HLM model because of the much larger number of parameters generated for each group, which is unlike the CRC model in that these are also the main quantities of interest.

Alternative approaches to the problem are to either make this calculation either more simple or more complex. The more simple approaches would adapt various OLS methods selectively to reduce the amount of data. For example, selected variables could be reduced into the form of categorical data, and least squares regression could be carried out using group-level predictors and interactions. However, as noted previously, this leads to potential problems with multicollinearity and does not allow individual observations to interact with group-level predictors, removing some of the flexibility of the HLM approach.

The more complex approaches to increasing the computational cost generally require rewriting the Gibbs sampler for speed. Writing a Gibbs sampler by hand would likely result in significant speed improvements over the JAGS module. Parts of the Gibbs sampling process can be parallelized, for example, the iteration of chains from multiple initial starting points, although parallelizing the MCMC process used to achieve convergence in the actual calculation of an individual chain introduces complex timing issues (Wilkinson, 2006).

6.7 Tables & Figures

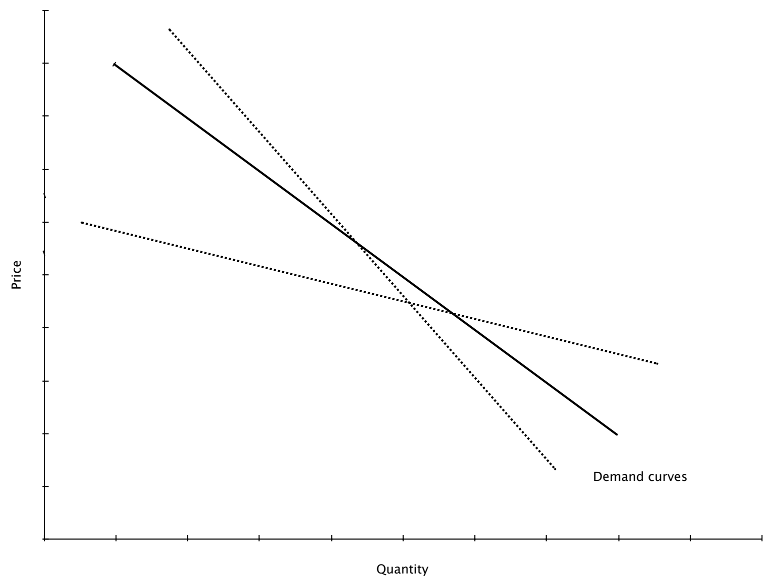

Figure 6.1: Various demand curves from estimated parameters

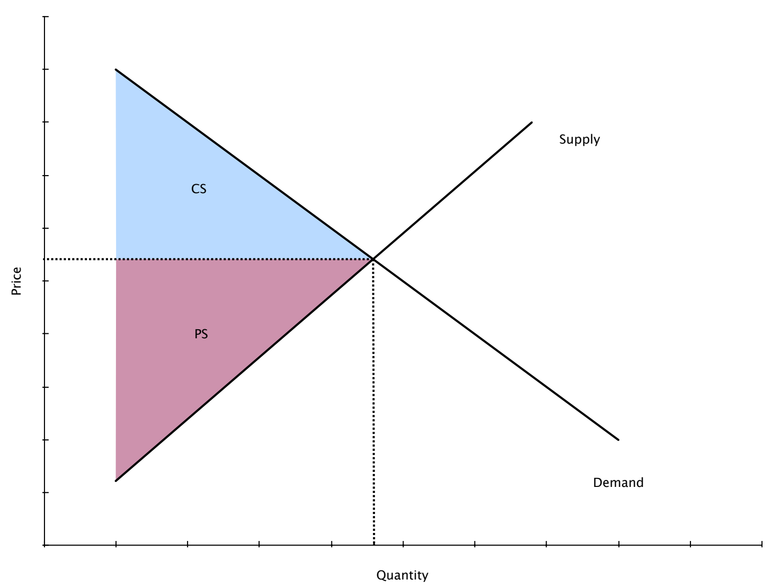

Figure 6.2: Consumer and producer surplus

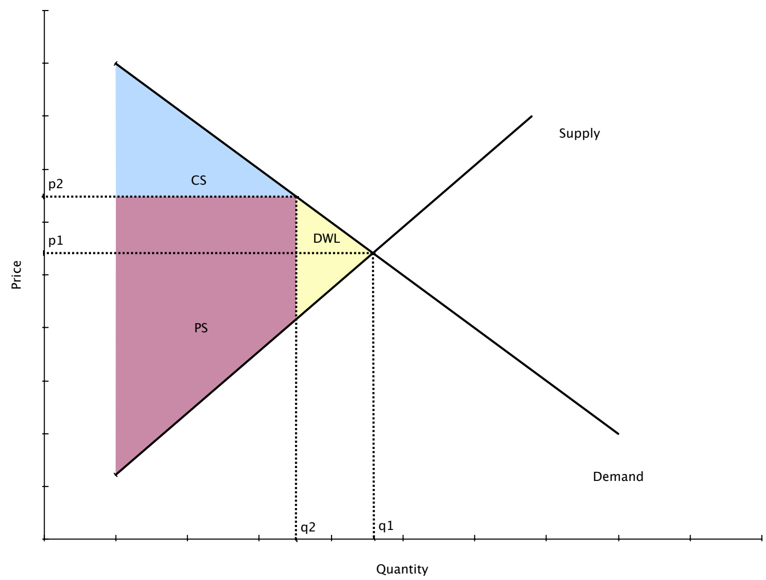

Figure 6.3: Change in Marshallian consumer surplus due to price change

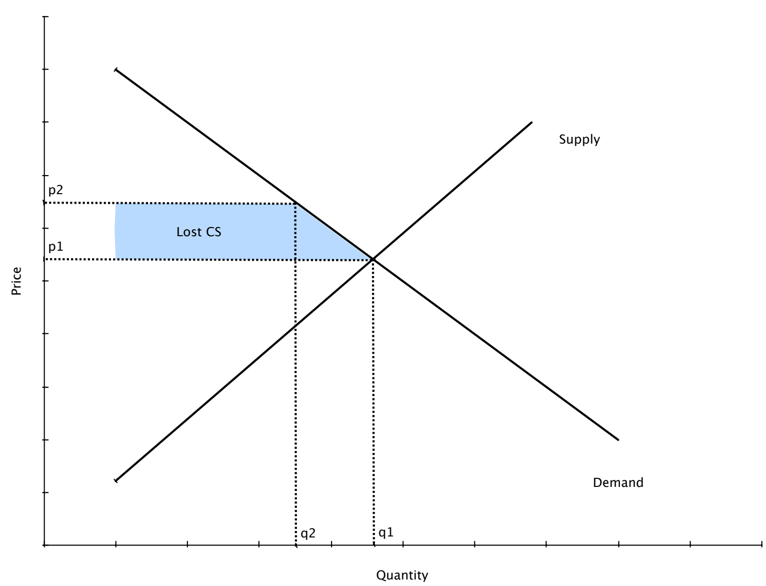

Figure 6.4: Lost Marshallian consumer surplus due to price change

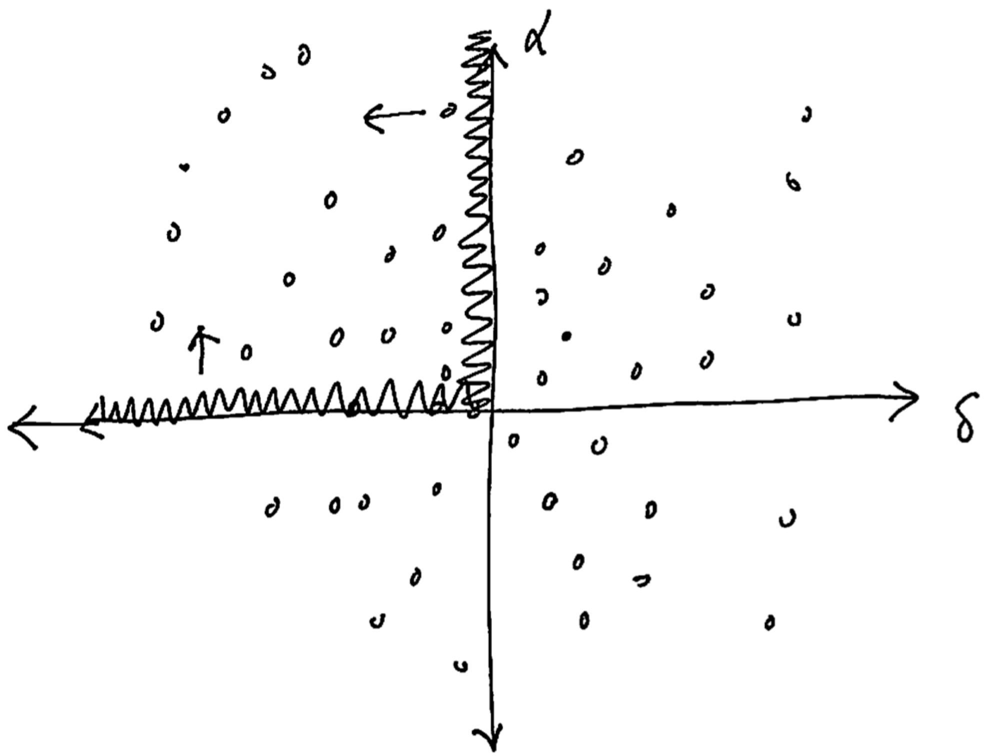

Figure 6.5: Restricted Quadrant for Parameter Values. Both economic and statistical theory place restrictions on the estimated parameter values. In the parameter space for α and δ, either an improper Bayesian prior or a Gamma distribution can be used to restrict them to the fourth quadrant.

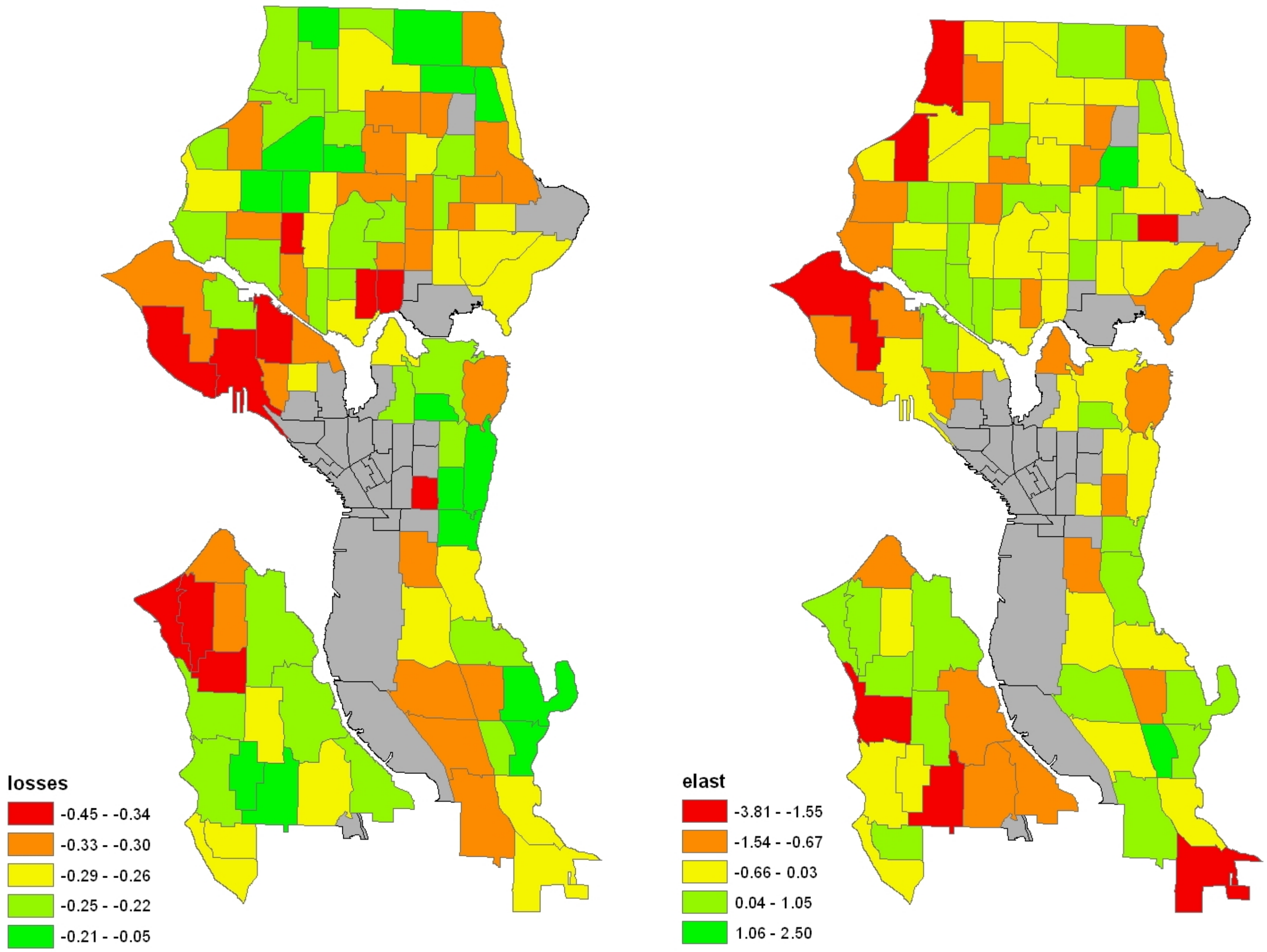

Figure 6.6: Consumer surplus losses as percentage of total bill. The left hand side shows the Marshallian CS losses as a percentage of the total bill by tract, and the right hand side shows the calculated elasticities by tract. Both maps are based on results from model CRC-2T from Chapter 4.

			Acct.-level models			Tract-level models		
Source	Tier	Units	A1	A3	A6	T1	T3	T6
Obs. average monthly bill	1,2	TB ($)	$ 23.54	$ 23.78	$ 24.56	$ 22.14	$ 22.35	$ 22.88
	3	TB ($)	$ 150.64	$ 153.42	$ 150.96	$ 138.53	$ 152.21	$ 149.82
Simulated welfare and errors	1,2	CV ($)	$ 2.14	$ (3.66)	$ (3.66)	$ 0.31	$ (1.22)	$ (8.85)
(Adamowicz et al., 1989)		sd	(152.20)	(168.06)	(201.91)	(69.24)	(88.76)	(87.54)
		% of TB	9%	-15%	-15%	1%	-5%	-39%
	3	CV ($)	$ 29.89	$ (7.32)	$ (8.24)	$ 64.05	$ 0.31	$ (17.69)
		sd	(954.04)	(914.39)	(1,217.26)	(146.40)	(960.75)	(1,126.67)
		% of TB	20%	-5%	-5%	46%	0%	-12%
Aberrant bootstrapping	1,2	CV ($)	$ (0.14)	$ 0.01	$ 0.14	$ (0.09)	$ 0.03	$ (0.04)
(Kling and Sexton, 1990)		sd	(11.71)	(11.69)	(11.32)	(11.08)	(11.95)	(12.08)
		% of TB	-0.7%	0.0%	0.7%	-0.4%	0.1%	-0.2%
	3	CV ($)	$ 30.39	$ 13.93	$ (0.87)	$ 55.51	$ 10.42	$ 7.07
		sd	(27.61)	(42.66)	(38.89)	(30.43)	(71.66)	(69.97)
		% of TB	22.1%	10.1%	-0.6%	40.3%	7.6%	5.1%

Table 6.1: Calculated Average CV Broken Out by Tier. Based on predicted consumption and estimated parameters from chapter 5. Although the aberrant bootstrapping method suggested by Kling and Sexton (1990) significantly reduces the uncertainty associated with the estimated welfare measures, the errors are still much larger than the estimated welfare effects.

		Acct.-level models			Tract-level models		
Source	Units	A1	A3	A6	T1	T3	T6
Obs. average monthly bill	TB ($)	$ 24.47	$ 24.47	$ 24.47	$ 24.98	$ 24.98	$ 24.98
Simulated welfare and errors	Average CV ($)	$ 1,575.34	$ (55.48)	$ 170.65	$ (413.99)	$ (53.23)	$ (237.11)
(Adamowicz et al., 1989)	Average sd	(75,471.93)	(17,536.28)	(16,497.04)	(18,745.03)	(5,276.59)	(14,878.54)
	25% quintile	-122.80	-142.88	-110.30	-105.98	-56.48	-118.72
	75% quintile	100.13	146.20	128.51	80.70	55.32	117.42
Aberrant bootstrapping	Average CV ($)	$ (0.44)	$ (0.21)	$ (1.13)	$ (0.27)	$ (0.52)	$ (0.02)
(Kling and Sexton, 1990)	Average sd	(13.03)	(13.49)	(13.00)	(14.03)	(13.59)	(14.21)
	25% quintile	-10.84	-11.30	-11.83	-11.97	-11.68	-12.17
	75% quintile	10.01	10.87	9.34	11.47	10.63	11.81

Table 6.2: Calculated Average CV for all Groups. Based on predicted consumption and estimated parameters from chapter 5. Although the aberrant bootstrapping method suggested by Kling and Sexton (1990) significantly reduces the uncertainty associated with the estimated welfare measures, the errors are still typically much larger than the estimated welfare effects for most groups (either by account or census tract).

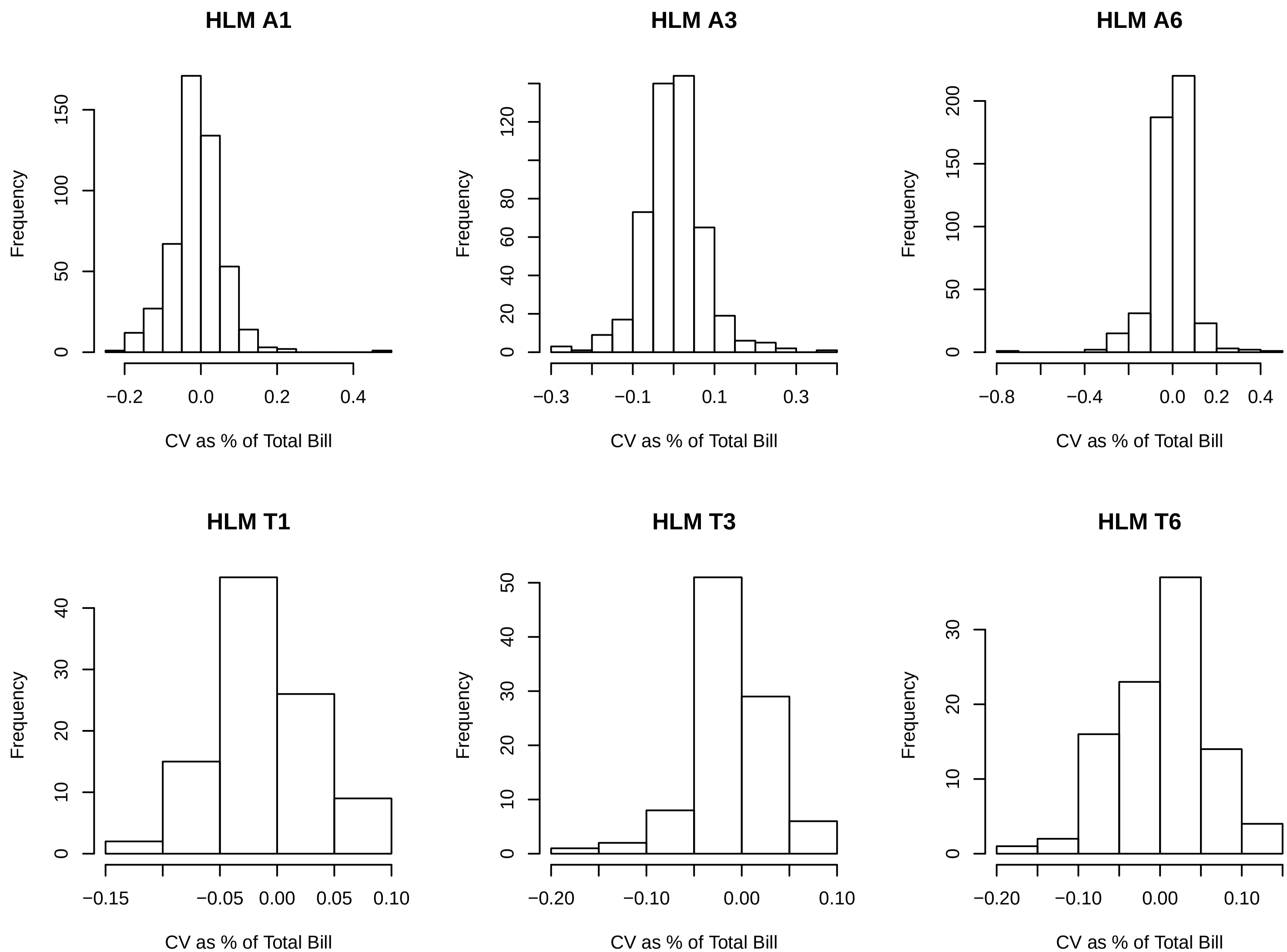

Figure 6.7: Distribution of Compensating Variation (CV) as a Percentage of Total Bill. Uses predicted from HLM models developed in Chapter 5. Total bill is defined as the average monthly water bill by account and tract.

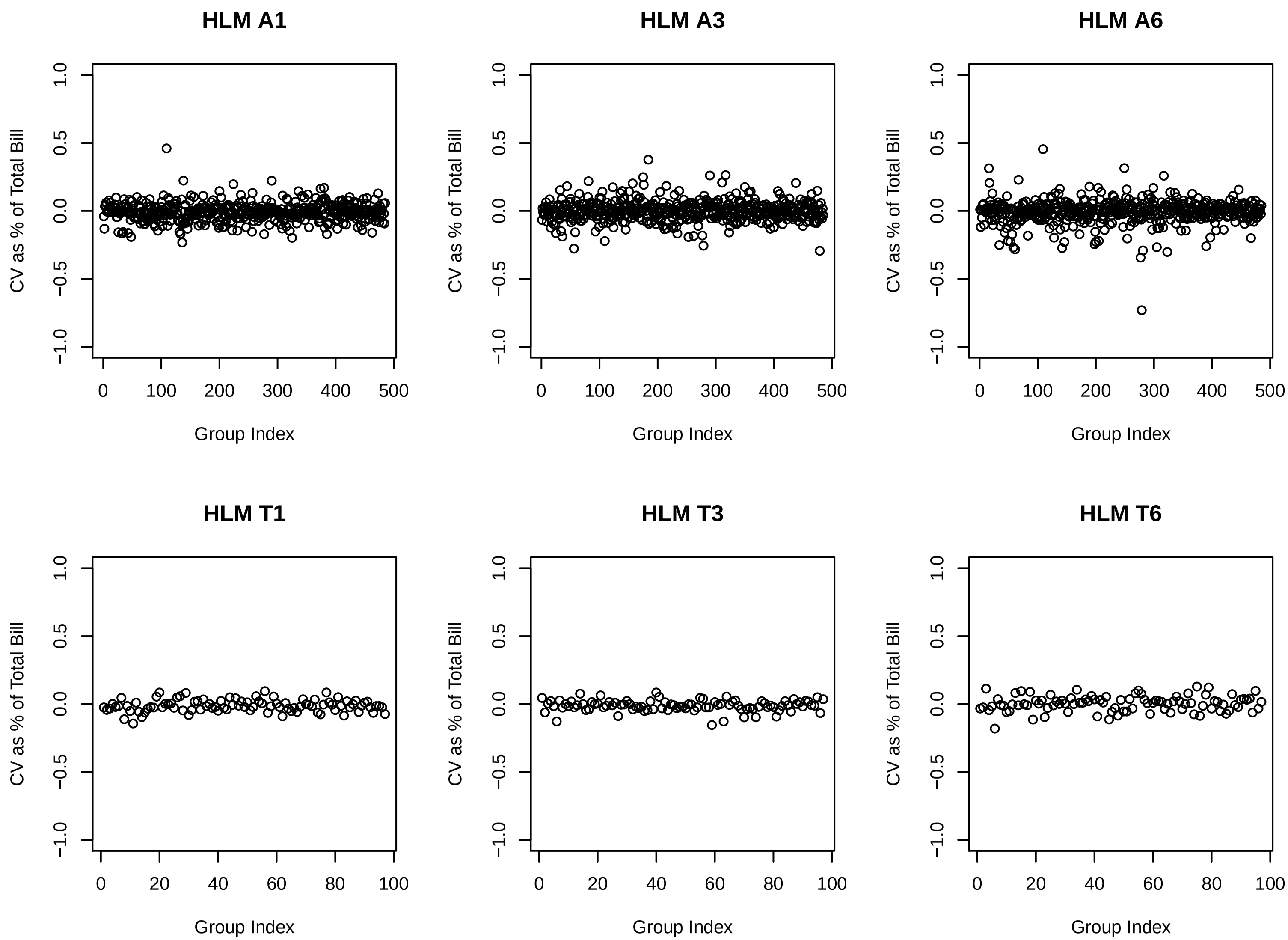

Figure 6.8: Distribution of Compensating Variation (CV) as a Percentage of Total Bill. Uses predicted from HLM models developed in Chapter 5. Total bill is defined as the average monthly water bill by account and tract. Error bars are not shown, because unfortunately they are typically more than two orders of magnitude larger than the calculated welfare measures, and would therefore obscure them.

Chapter 7

CONCLUSIONS & FUTURE WORK

Chapter 1 of this dissertation introduced the main policy of interest, a pricing policy change introduced in the City of Seattle in 2001. The research question asked was, how did this change in pricing policy affect water consumption for different users? The remainder for the introduction then argued why prices have emerged as a key policy of interest; why an integrated approach to the issues of simultaneity, micro-data, and heterogeneity is a necessary and useful step forward; and how this would build on several emerging research trends. Chapter 2 summarized the previous approaches in the literature, with a particular emphasis on how data limitations have shaped the subsequent literature, and Chapter 3 described the unique new dataset used for the subsequent analytical chapters.

Chapter 4 introduced the correlated random coefficients (CRC) model from labor economics, in order to allow varying coefficients on the marginal price instrument, which describes how different users to respond differently to price changes. This flexible model was then applied to different geographic groupings, taking advantage of the available disaggregated dataset. Not only were the random effect (RE) and CRC models found to be more accurate and flexible than other pooled 2SLS methods, the elasticities calculated are consistent with the range of estimates from previous meta-analyses.

Chapter 5 introduced a hierarchical linear model (HLM) model for water demand. Building on the groupings used in Chapter 4, the CRC model is shown simply to be a special case of the HLM model. The HLM model is then extended from these results to explain the sources of variation among households and census tracts. For households, or account-level models, the expected relationships between house value and lot sizes, and the price elasticity, are observed. Similarly, for the tract-level models, the expected relationships between median income and median house value are also observed, although the smaller number of groups leads to weaker results. However, because group-level results can be related to group-level information, it is now possible to draw policy-relevant, geographically-specific conclusions for particular census tracts. At the end of Chapter 5, the local average treatment effect (LATE) was calculated at the tract-level, but the HLM results were found to result in insignificant or implausible results.

Chapter 6 explored different ways to measure the impact of the pricing policy change. In addition to the LATE effect calculated at the end of Chapter 5, I performed additional calculations for the impact to Marshallian consumer surplus (CS) and an exact calculation of consumer surplus, introduced by Hausman (1981). Based on the CRC results, it is possible to obtain relatively clear results for the change in Marshallian CS as a result of the policy change. These can also be aggregated by census tract and mapped, and in Figure 6.6, the losses by census tract are calculated and found to range from 5-45% of the consumer's average total bills. Based on the HLM results, it is difficult to obtain welfare measures from the exact calculation with well-behaved statistical properties. Ultimately, these results need to be compared to the benefit of the avoided costs of new supplies, in order to obtain a full benefit-cost analysis of the impact of this conservation policy.

Based on the results of this dissertation, I recognize that there are ample opportunities to extend and improve on this work in the future. These dissertation results could be greatly extended by improving the methods in three main areas. First, the size and speed of the calculations used to obtain these results is a major problem, enough to limit the amount of data that can be used effectively. In Chapter 4, the CRC analysis only uses 1% of the available data, and in Chapter 5, because Markov Chain Monte Carlo (MCMC) calculations are particularly slow, this analysis in turn only uses 5% of the CRC data. Although it is unusual for statistical methods to be limited by speed and not data, particularly in the social sciences, this will increasingly become a problem as ubiquitous sensor data becomes increasingly available. Second, given the complexity of the model structures, there is a clear need for comparative metrics for model performance and fit. Calculation of explained variance at each level in the HLM models would aid in assessing the relative performance and fit of various models. Cross-validation and root mean square error (RMSE) seems to be the best way to measure the predictive accuracy of the models. Third, and finally, the estimated parameters obtained in Chapter 5 are not accurate enough to calculate the more complex welfare and policy measures in Chapter 6, such as the LATE or exact CS measure.

For future work, based on these recognized limitations, I therefore intend to compare the predictive accuracy of the models using RMSE on out-of-sample tests to perform cross-validation, and to test the grouping effectiveness. In addition, improving MCMC speed could lead to greatly enhanced results. In particular, parallel approaches to the statistical calculations, particularly in the Gibbs sampling process, could lead to faster calculations and potentially more accurate results. Image analysis techniques offer two possible approaches that may allow faster processing of large datasets, such as Markov random fields or simulated annealing methods. Finally, the analysis in this dissertation was intended to compare and contrast the results from different groupings of disaggregated data, such as by account *or* by tract. Other than the speed of the calculations, there is no reason why crossed-effect models using multiple groupings and datasets could not be used to obtain improved predictive models.

Another possible avenue for future work is to focus on the causal inference of the effects of the pricing policy. Although this model has sought to develop models for predictive purposes, it would also be of interest to assess the impacts of the pricing policy using methods of causal inference, to see what change in water consumption is directly attributable to the pricing policy. This will require renewed attention to the research design and character of the observational data.

Chapter 8

BIBLIOGRAPHY

BIBLIOGRAPHY

Adamowicz, W. L., J. J. Fletcher, and T. Graham-Tomasi (1989, May). Functional form and the statistical properties of welfare measures. *American Journal of Agricultural Economics 71*(2), 414–421.

Agthe, D. E., R. B. Billings, J. L. Dobra, and K. Raffiee (1986). Simultaneous equation demand model for block rates. *Water Resources Research WRERAO 22*(1).

Angrist, J. D. and J. S. Pischke (2009). *Mostly harmless econometrics: an empiricist's companion*. Princeton Univ Pr.

Arbues, F., M. Garcia-Valinas, and R. Martinez-Espineira (2003, March). Estimation of residential water demand: a state-of-the-art review. *Journal of Socio-Economics 32*(1), 81–102.

Arbus, F., R. Barbern, and I. Villana (2004). Price impact on urban residential water demand: a dynamic panel data approach. *Water Resources Research 40*(11), W11402.

Barnett, T. P., J. C. Adam, and D. P. Lettenmaier (2005). Potential impacts of a warming climate on water availability in snow-dominated regions. *Nature 438*(7066), 303309.

Bates, D. (2005). Fitting linear models in r using the lme4 package. *R News 5*(1), 27–30.

Bates, D. (2010). lme4: Mixed-Effects modeling with r.

Berry, B. (1990). Urbanization. In B. L. Turner (Ed.), *The Earth as Transformed by Human Action*, pp. 103–119. New York: Cambridge University Press.

Billings, R. B. (1982). Specification of block rate price variables in demand models. *Land Economics 58*(3), 386–394.

Billings, R. B. and D. E. Agthe (1980). Price elasticities for water: A case of increasing block rates. *Land Economics 56*(1), 73–84.

Blaug, M. (1996). *Economic Theory in Retrospect*. Cambridge University Press.

Blumstein, C., B. Krieg, L. Schipper, and C. York (1980). Overcoming social and institutional barriers to energy conservation. *Energy 5*(4), 355–372.

Bockstael, N. E., W. M. Hanemann, and E. Strand (1989). Measuring the benefits of water quality improvements using recreation demand models. volume 2 of benefits analysis using indirect or imputed market methods.

Bockstael, N. E. and I. E. Strand (1987, February). The effect of common sources of regression error on benefit estimates. *Land Economics 63*(1), 11–20.

Card, D. (2001, September). Estimating the return to schooling: Progress on some persistent econometric problems. *Econometrica 69*(5), 1127–1160. ArticleType: primary_article / Full publication date: Sep., 2001 / Copyright 2001 The Econometric Society.

Casella, G., R. L. Berger, and R. L. Berger (2002). *Statistical inference*. Duxbury Pacific Grove, CA.

Chicoine, D. L. and G. Ramamurthy (1986). Evidence on the specification of price in the study of domestic water demand. *Land Economics 62*(1), 26–32.

Coleman, E. A. (2009). A comparison of Demand-Side management strategies using disaggregate data. *Public Works Management Policy 13*(3), 215–223.

da Conceicao Cunha, M. and J. Sousa (1999). Water distribution network design optimization: simulated annealing approach. *Journal of Water Resources Planning and Management 125*, 215.

Dalhuisen, J. M., R. Florax, H. L. F. de Groot, and P. Nijkamp (2003). Price and income elasticities of residential water demand: A Meta-Analysis. *Land Economics 79*(2), 292–308.

Danielson, L. E. (1979). An analysis of residential demand for water using micro Time-Series data. *Water Resources Research 15*(4), 763–767.

Deller, S. C., D. L. Chicoine, and G. Ramamurthy (1986). Instrumental variables approach to rural water service demand. *Southern Economic Journal 53*(2), 333–346.

Deoreo, W. B., A. Dietemann, T. Skeel, P. W. Mayer, D. M. Lewis, and J. Smith (2001). Retrofit realities. *Journal American Water Works Association 93*(3), 5872.

Dryzek, J. S. (2005). *The Politics of the Earth: Environmental Discourses.* Oxford: Oxford University Press.

Duhigg, C. (2009a, December). Millions in U.S. drink contaminated water, records show - series - NYTimes.com. *New York Times*.

Duhigg, C. (2009b, September). Toxic waters - clean water laws are neglected, at a cost in suffering - series - NYTimes.com. *New York Times*.

Duhigg, C. (2010, March). Saving U.S. water and sewer systems would be costly. *The New York Times*.

Economy, E. (2004). *The river runs black: the environmental challenge to China's future.* Cornell Univ Pr.

Espey, M., J. Espey, and W. D. Shaw (1997). Price elasticity of residential demand for water: A meta-analysis. *Water Resources Research 33*(6), 1369–1374.

Eusuff, M. M. and K. E. Lansey (2003, May). Optimization of water distribution network design using the shuffled frog leaping algorithm. *Journal of Water Resources Planning and Management 129*(3), 210–225.

Fisher, W. D. (1969). *Clustering and Aggregation in Economics.* Baltimore: Johns Hopkins Press.

Flory, B. (2008, October). Personal communication.

Foster, H. S. and B. R. Beattie (1979). Urban residential demand for water in the united states. *Land Economics 55*(1), 43–58.

Foster, H. S. and B. R. Beattie (1981, May). Urban residential demand for water in the united states: Reply. *Land Economics 57*(2), 257–265. ArticleType: primary_article / Full publication date: May, 1981 / Copyright 1981 The Board of Regents of the University of Wisconsin System.

Fox, J. (1979). Simultaneous equation models and two-stage least squares. *Sociological methodology 10*, 130150.

Fox, J. (2006). Structural equation modeling with the sem package in r. *Structural Equation Modeling 13*(3), 465–486.

Freeman, M. (2003). *The Measurement of Environmental and Resource Values: Theory and Methods.* RFF Press (Resources for the Future).

Garen, J. (1984, September). The returns to schooling: A selectivity bias approach with a continuous choice variable. *Econometrica 52*(5), 1199–1218.

Gelman, A. (2006). Multilevel (Hierarchical) modeling: what it can and cannot do. *Technometrics 48*(3), 432435.

Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin (2004). *Bayesian Data Analysis.* Boca Raton: Chapman & Hall / CRC.

Gelman, A. and J. Hill (2007). *Data analysis using regression and multilevel/hierarchical models.* New York, NY: Cambridge University Press.

Gelman, A. and I. Pardoe (2006). Bayesian measures of explained variance and pooling in multilevel (hierarchical) models. *Technometrics 48*(2), 241251.

Gertner, J. (2007, October). The future is drying up. *The New York Times*.

Green, H. A. J. (1964). *Aggregation in Economic Analysis: An Introductory Survey.* Princeton University Press.

Griffin, A. H., W. E. Martin, and J. C. Wade (1981, May). Urban residential demand for water in the united states: Comment. *Land Economics 57*(2), 252–256.

Grimond, J. (2010, May). For want of a drink: a special report on water. *Economist 395*(8683), 1–20.

Hanemann, W. M. (1998). Determinants of urban water use. In D. D. Baumann, J. J. Boland, and W. M. Hanemann (Eds.), *Urban water demand management and planning.* McGraw-Hill Corporation.

Hanemann, W. M. and C. Nauges (2005). Heterogeneous responses to water conservation programs: The case of residential users in los angeles. Number 1026.

Hausman, J. A. (1981, September). Exact consumer's surplus and deadweight loss. *The American Economic Review 71*(4), 662–676.

Heckman, J. and E. Vytlacil (1998). Instrumental variables methods for the correlated random coefficient model: Estimating the average rate of return to schooling when the return is correlated with schooling. *The Journal of Human Resources 33*(4), 974–987.

Heckman, J. J. (2001). Micro data, heterogeneity, and the evaluation of public policy: Nobel lecture. *Journal of Political Economy 109*(4), 673–748.

Hewitt, J. A. (2000). An investigation into the reasons why water utilities choose particular residential rate structures. *The political economy of water pricing reforms*, 259–277.

Hewitt, J. A. and W. M. Hanemann (1995). A Discrete/Continuous choice approach to residential water demand under block rate pricing. *Land Economics 71*(2), 173–192.

Houthakker, H. S. (1951). Some calculations of electricity consumption in great britain. *Journal of the Royal Statistical Society, Series A 114*, 351–71.

Howe, C. W. and F. P. Linaweaver (1967). The impact of price on residential water demand and its relation to system design and price structure. *Water Resources Research 3*(1), 13–32.

Hoyle, B. (2008). The energy-water nexus: deja-vu all over again? *Nature Reports Climate Change* (0804).

Johns, A. (2010, April). Water hits boiling point. *Chattanooga Times Free Press*.

Jones, C. V. and J. R. Morris (1984). Instrumental price estimates and residential water demand. *Water Resources Research 20*(2), 197–202.

Karl, T. R., C. N. Williams, P. J. Young, and W. M. Wendland (1986, February). A model to estimate the time of observation bias associated with monthly mean maximum, minimum and mean temperatures for the united states. *Journal of Applied Meteorology 25*(2), 145–160.

Kenny, J., N. Barber, S. Hutson, K. Linsey, J. Lovelace, and M. Maupin (2009, October). Estimated use of water in the united states in 2005: U.S. geological survey circular 1344. U.S. Geological Survey Circular 1344, U.S. Geological Survey.

Kling, C. L. (1991, November). Estimating the precision of welfare measures. *Journal of Environmental Economics and Management 21*(3), 244–259.

Kling, C. L. and R. J. Sexton (1990, May). Bootstrapping in applied welfare analysis. *American Journal of Agricultural Economics 72*(2), 406–418.

Knowles, N. and D. R. Cayan (2002). Potential effects of global warming on the Sacramento/San joaquin watershed and the san francisco estuary. *Geophysical Research Letters 29*(18), 381.

Koppelman, F. S. (1974). Prediction with disaggregate data: The aggregation issue. Transportation Research Record 527, Transportation Research Board.

Kousky, C. and S. Schneider (2003). Global climate policy: Will cities lead the way? *Climate Policy*.

Krinsky, I. and A. L. Robb (1986, November). On approximating the statistical properties of elasticities. *The Review of Economics and Statistics 68*(4), 715–719.

Laird, N. M. and J. H. Ware (1982, December). Random-Effects models for longitudinal data. *Biometrics 38*(4), 963–974.

Layton, D. F. and G. Brown (2000). Heterogeneous preferences regarding global climate change. *The Review of Economics and Statistics 82*(4), 616624.

Lee, P. M. (2004). *Bayesian statistics* (Third ed.). Arnold London, UK:.

Lutzenhiser, L., B. Hackett, and H. Schutz (1987). Social variation and electricity consumption in san diego, california: Exploratory data analysis and the california energy commissions electricity demand forecasting model. *UER-196. University of California Energy Research Group, Berkeley, California*.

Mcfadden, D. (2001). Economic choices. *American Economic Review 91*(3), 351–378.

Morgan, S. L. and C. Winship (2007). *Counterfactuals and causal inference: Methods and principles for social research*. Cambridge Univ Pr.

Murdock, S. H., D. E. Albrecht, R. R. Hamm, and K. Backman (1991). Role of sociodemographic characteristics in projections of water use. *Journal of Water Resources Planning and Management 117*(2), 235–251.

Nataraj, S. and W. M. Hanemann (2008, November). Does marginal prices matter? a regression discontinuity approach to estimating water demand.

Nations, U. (1992). Agenda 21: The united nations programme of action from rio. Technical report, United Nations.

Nieswiadomy, M. L. and D. J. Molina (1989). Comparing residential water demand estimates under decreasing and increasing block rates using household data. *Land Economics 65*(3), 280–289.

Nieswiadomy, M. L. and D. J. Molina (1991). A note on price perception in water demand models. *Land Economics 67*(3), 352–359.

Nordin, J. A. (1976). A proposed modification of taylor's demand analysis: comment. *The Bell Journal of Economics 7*, 719–721.

Olmstead, S. M., W. M. Hanemann, and R. N. Stavins (2007). Water demand under alternative price structures. *Journal of Environmental Economics and Management 54*(2), 181–198.

Olmstead, S. M. and R. Stavins (2008). *Comparing price and non-price approaches to urban water conservation*. National Bureau of Economic Research Cambridge, Mass., USA.

Opaluch, J. J. (1982). Urban residential demand for water in the united states: Further discussion. *Land Economics 58*(2), 225–227.

Opaluch, J. J. (1984). A test of consumer demand response to water prices: reply. *Land Economics 60*(4), 417–421.

Pearl, J. (2000). *Causality: models, reasoning, and inference*. Cambridge Univ Pr.

Peet, J. (2003). Priceless: A survey of water. *The Economist 368*.

Pint, E. M. (1999). Household responses to increased water rates during the california drought. *Land Economics 75*(2), 246–266.

Plummer, M. (2010). JAGS - just another gibbs sampler. http://www-fis.iarc.fr/~martyn/software/jags/.

Polebitski, A. S. and R. N. Palmer (2009). Seasonal residential water demand forecasting for census tracts. *Journal of Water Resources Planning and Management 1*, 3.

Portney, K. E. (2003). *Taking Sustainable Cities Seriously: Economic Development, the Environment, and Quality of Life in American Cities (American and Comparative Environmental Policy)*. The MIT Press.

Postel, S. L. (2000). Entering an era of water scarcity: the challenges ahead. *Ecological Applications 10*(4), 941–948.

Postel, S. L., G. C. Daily, and P. R. Ehrlich (1996). Human appropriation of renewable fresh water. *Science 271*(5250), 785.

Rees, W. (1992). Ecological footprints and appropriated carrying capacity: What urban economics leaves out. *Environment and Urbanization 4*, 121–130.

Renwick, M. E. and R. D. Green (2000). Do residential water demand side management policies measure up? an analysis of eight california water agencies. *Journal of Environmental Economics and Management 40*(1), 37–55.

Rogers, P., R. de Silva, and R. Bhatia (2002). Water is an economic good: How to use prices to promote equity, efficiency, and sustainability. *Water Policy 4*, 117.

Rossi, P. H., M. W. Lipsey, and H. E. Freeman (2004). *Evaluation: A Systematic Approach*. Sage Publications.

Savic, D. A. and G. A. Walters (1997, March). Genetic algorithms for Least-Cost design of water distribution networks. *Journal of Water Resources Planning and Management 123*(2), 67–77.

Schefter, J. E. and E. L. David (1985, August). Estimating residential water demand under Multi-Part tariffs using aggregate data. *Land Economics 61*(3), 272–280.

Schneider, M. L. and E. E. Whitlatch (1991). User-Specific water demand elasticities. *Journal of Water Resources Planning and Management 117*(1), 52–73.

Shin, J. S. (1985). Perception of price when price information is costly: evidence from residential electricity demand. *The Review of Economics and Statistics 67*(4), 591–598.

Silberberg, E. and W. Suen (2001). *The structure of economics: a mathematical analysis* (3rd ed.). New York: McGraw-Hill.

Slesnick, D. T. (1998, December). Empirical approaches to the measurement of welfare. *Journal of Economic Literature 36*(4), 2108–2165.

Slesnick, D. T. (2008). Consumer surplus.

Smil, V. (2003). *The Earth's biosphere: Evolution, dynamics, and change*. The MIT Press.

Specter, M. (2006). The last drop. *The New Yorker*, 5071.

Stauffer, D. (2004). Emerging water shortages are no mirage. *Planning Commissioners Journal* (54), 4.

Su, Y. and M. Yajima (2010). CRAN - package r2jags. http://cran.at.r-project.org/web/packages/R2jags/.

Taylor, L. D. (1975). The demand for electricity: A survey. *The Bell Journal of Economics 6*(1), 74–110.

Team, R. D. C. (2009). R: A language and environment for statistical computing.

Terza, J. V. (1986). Determinants of household electricity demand: A Two-Stage probit approach. *Southern Economic Journal 52*(4), 11311139.

Theil, H. (1954). *Linear aggregation of econometric relations*. Amsterdam, North-Holland.

Timmins, C. (2003). Demand-Side technology standards under inefficient pricing regimes. *Environmental and Resource Economics 26*(1), 107–124.

Train, K. (1991). *Optimal regulation: the economic theory of natural monopoly*. MIT Press.

Vitousek, P. M., H. A. Mooney, J. Luchenco, and J. M. Melillo (1997). Human domination of the earth's ecosystems. *Science 277*.

Watkins, K., L. Carvajal, D. Coppard, R. Fuentes, A. Ghosh, C. Giamberardini, C. Johansson, P. Seck, C. Ugaz, and S. Yaqub (2006). Human development report 2006. *Beyond Scarcity: Power, Poverty and the Global Water Crisis. United Nations Development Programme (UNDP), New York*.

Webber, M. E. (2008, October). Energy versus water: Solving both crises together. *Scientific American*.

Weimer, D. (2008). Cost-benefit analysis.

Wilder, R. P. and J. F. Willenborg (1975). Residential demand for electricity: a consumer panel approach. *Southern Economic Journal 42*(2), 212217.

Wilkinson, D. J. (2006). Parallel bayesian computation. *STATISTICS TEXTBOOKS AND MONOGRAPHS 184*, 477.

Willig, R. D. (1976). Consumers surplus without apology. *American Economic Review 66*(4), 589–597.

Wooldridge, J. M. (1997, October). On two stage least squares estimation of the average treatment effect in a random coefficient model. *Economics Letters 56*(2), 129–133.

Wooldridge, J. M. (2003, May). Further results on instrumental variables estimation of average treatment effects in the correlated random coefficient model. *Economics Letters 79*(2), 185–191.

Yardley, J. (2004, September). The new uprooted: In a tidal wave, china's masses pour from farm to city. *New York Times*.

Appendix A

BUGS CODE FOR GIBBS SAMPLING

```
model {
for (i in 1:n){
y[i] ~ dnorm(y.hat[i], tau.y)
y.hat[i] <- inprod(b.0[], X.0[i,]) + inprod(B[group[i],], X[i,])
pred.y[i] ~ dnorm(y.hat[i], tau.y)
y.hat0[i] <- inprod(b.0[], X.0[i,]) + inprod(B[group[i],], X.p0[i,])
pred.y0[i] ~ dnorm(y.hat0[i], tau.y)
}
tau.y <- pow (sigma.y, -2)
sigma.y ~ dunif (0, 100)
for (k in 1:K.0){
b.0[k] ~ dnorm(0, 1e-4)
}
for (k in 1:K){
for (j in 1:J){
B[j,k] <- B.raw[j,k]
}
}
for (j in 1:J){
B.raw[j, 1:K] ~ dmnorm(B.raw.hat[j,], Tau.B.raw[,])
for (k in 1:K){
B.raw.hat[j,k] <- inprod(G.raw[k,], U[j,])
}
}
```

```
for (k in 1:K){
for (l in 1:L){
G[k,l] <- G.raw[k,l]
G.raw[k,l] ~ dnorm(0, 1e-4)
}
}
Tau.B.raw[1:K, 1:K] ~ dwish (W[,], df)
df <- K + 1
Sigma.B.raw[1:K, 1:K] <- inverse(Tau.B.raw[,])
for (k in 1:K){
for (k.prime in 1:K){
rho.B[k, k.prime] <- Sigma.B.raw[k, k.prime] /
sqrt(Sigma.B.raw[k,k]*Sigma.B.raw[k.prime, k.prime])
}
sigma.B[k] <- sqrt(Sigma.B.raw[k,k])
}
}
```

Appendix A

DERIVATION OF RATIO OF TWO STANDARD NORMALS

Casella et al. (2002) offers a fuller version of this derivation on pages 156-162.

For a bivariate random vector (X, Y) with a known probability distribution, we can define a new bivariate random vector (U, V) defined by transformations $U = g_1(X, Y)$ and $V = g_2(X, Y)$, where $g_1(x, y)$ and $g_2(x, y)$ are functions. If B belongs to the set of real numbers, then $(U, V) \in B$ and if and only if $(X, Y) \in A$ where entire set A belongs to the set B, then the probability distribution of (U, V) can be completely determined by the probability distribution of (X, Y). The joint pdf of (U, V), $f_{U,V}(u, v)$ can therefore be computed from the joint pdf of (X,Y) by

$$f_{U,V}(u, v) = P(U = u, V = v) = P((X, Y) \in A) = \sum_{(x,y) \in A_{uv})} f_{X,Y}(x, y)$$

For a one-to-one transformation where each $(u, v) \in B$ maps to only one $(x, y) \in A$, we then solve the equations $u = g_1(x, y)$ and $v = g_2(x, y)$ for x, y in terms of u, v. Denoting these inverse transformations $u = h_1(x, y)$ and $v = h_2(x, y)$ we can calculate the Jacobian of the transformation,

$$J = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \frac{\partial x}{\partial u}\frac{\partial y}{\partial v} - \frac{\partial y}{\partial u}\frac{\partial x}{\partial v}$$

where

$$\frac{\partial x}{\partial u} = \frac{\partial h_1(u, v)}{\partial u}, \quad \frac{\partial x}{\partial v} = \frac{\partial h_1(u, v)}{\partial v}, \quad \frac{\partial y}{\partial u} = \frac{\partial h_2(u, v)}{\partial u}, \quad \frac{\partial y}{\partial v} = \frac{\partial h_2(u, v)}{\partial v}$$

Therefore the joint pdf $f_{U,V}(u, v)$ and $f_{X,Y}(x, y)$ are related by:

$$f_{U,V}(u, v) = f_{X,Y}(h_1(u, v), h_2(u, v)) \times |J|$$

We can now specify the bivariate transformation $U = X/Y$ and $V = |Y|$. This restriction of V to the absolute value of Y is critical in order to preserve a one-to-one transformation, since the pairs (x, y) and $(-x, -y)$ would map to the same point (u, v). This can also be considered as necessary to preserve the identifiability of the any quantity that depends on x/y. Obtaining the inverse transformations $x = h_{11}(u, v) = uv$, $y = h_{21}(u, v) = v$, $x = h_{12}(u, v) = -uv$, and $y = h_{22}(u, v) = -v$, we get the Jacobians $J_1 = J_2 = v$. Using the joint bivariate normal distribution,

$$f_{X,Y}(x, y) = \frac{1}{2\pi} e^{-x^2/2} e^{-y^2/2}$$

and therefore

$$\begin{aligned} f_{U,V}(u, v) &= \frac{1}{2\pi} e^{-(uv)^2/2} e^{-v^2/2} |v| + \frac{1}{2\pi} e^{-(-uv)^2/2} e^{-(-v)^2/2} |v| \\ &= \frac{v}{\pi} e^{-(u^2+1)v^2/2}, \quad -\infty \leq u \leq \infty, \quad 0 \leq v \leq \infty \end{aligned}$$

The marginal pdf of U can be calculated, using a change of variable $z = v^2$, as

$$f_U(u) = \frac{1}{\pi(u^2+1)}, \infty \leq u \leq \infty$$

which is the pdf of a Cauchy random variable. Therefore, the ratio of two standard normals is distributed as a Cauchy random variable.

VITA

Yuin-Jen David Hsu previously worked in a number of fields, including urban planning, environmental design, structural engineering, real estate finance, and city government. He also holds degrees from Yale University, Cornell University, and the London School of Economics and Political Science. Before living in Seattle, he grew up in Amherst, Massachusetts, and in addition to studying in New Haven and Ithaca, he has lived in Boston, London, and New York City. In the summer of 2010, he will begin as an Assistant Professor in the Department of City and Regional Planning at the University of Pennsylvania, in Philadelphia.