

ABSTRACT

KUMINOFF, NICOLAI VLADIMIR. Recovering Preferences for Public Goods from a Dual-Market Locational Equilibrium. (Under the direction of V. Kerry Smith.)

This research extends the existing literature on revealed preference models of location choice by developing a “dual-market” sorting framework that uses information on households and their location choices in the housing and labor markets to infer their demand for local public goods. It also recognizes that households may differ in their relative preferences for those public goods. Four interrelated objectives are addressed. First, the analysis develops a theoretical model that relates households’ location choices in the housing and labor markets to the levels of local public goods provided by each of a finite set of communities. Second, it generalizes Epple and Sieg’s (1999) empirical model to recognize that: (i) households make a joint job-house choice and (ii) households differ in their job skills and in their relative preferences for different public goods. Third, the analysis uses the empirical model to develop a general equilibrium framework to simulate how households and markets would adjust to a large-scale change in the provision of a public good. Finally, the new dual-market estimation and simulation frameworks are used to analyze the welfare implications of a hypothetical air quality improvement that would allow the San Francisco-Sacramento region of Northern California to meet the state’s recently revised standards for ambient concentrations of ground-level ozone.

Recovering Preferences for Public Goods from a Dual-Market Locational Equilibrium

by

NICOLAI V. KUMINOFF

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

ECONOMICS

Raleigh

2006

APPROVED BY:

V. Kerry Smith
(Chair of the Advisory Committee)

Paul L. Fackler

Raymond B. Palmquist

Daniel J. Phaneuf

DEDICATION

To Joana

BIOGRAPHY

Nicolai V. Kuminoff was born in San Rafael, California on September 21st, 1977. After graduating from Terra Linda High School in 1995, he attended the University of California, Davis. While officially enrolled at UC Davis, Nick spent January through March of 1998 studying rainforest and coral reef ecosystems in Belize and Guatemala through the study abroad program offered by the Sierra Institute at UC Santa Cruz. Following his return, he took a position as an undergraduate research assistant at the University of California Agricultural Issues Center (AIC), where he continued to work while pursuing undergraduate and masters degrees at UC Davis. Nick received a Bachelor of Science degree in Agricultural Economics in 1999. The following year he received a Masters of Science degree in Agricultural Economics and accepted a new appointment as a research associate at AIC. He left AIC in August 2001 to pursue a doctorate in Economics at North Carolina State University. In September 2006, Nick joined the department of Agricultural and Applied Economics at Virginia Tech as an assistant professor.

ACKNOWLEDGEMENTS

I especially thank the chair of my committee, V. Kerry Smith, for his tireless efforts in guiding me through the dissertation—a tremendous investment of his time. I have been extremely fortunate to have had the opportunity to work with him. I also thank my other committee members: Paul Fackler, Ray Palmquist, and Dan Phaneuf. Through their research, teaching, and comments on my work, each committee member improved the quality of this dissertation and enhanced my appreciation of economics.

While at NC State, I also benefited from numerous discussions with Dave Butry, Carlos Carpio, Federico Ciliberto, Peyton Ferrier, “Uncle” Albert Pollatchek, Jaren Pope, Dan Sumner, Ada Wossink, and Chen Zhen. Funding for this research from the National Science Foundation is gratefully acknowledged.

TABLE OF CONTENTS

	Page
I. LIST OF TABLES.....	xi
II. LIST OF FIGURES.....	xiv
1. INTRODUCTION.....	1
2. REVEALED PREFERENCE MODELS OF LOCATION CHOICE.....	9
I. Introduction.....	10
II. Choice and Local Public Goods.....	11
III. Hedonic Models.....	13
(A) Property value model: continuous choice set.....	15
i. strategy #1: multiple markets.....	18
ii. strategy #2: restricting preferences.....	20
iii. strategy #3: fully specifying the utility function.....	22
iv. comparing the three strategies.....	24
(B) Property value model: discrete choice set.....	25
(C) Interregional hedonic model.....	28
(D) Summarizing the recent developments and remaining limitations.....	31
IV. Sorting Models.....	33
(A) Theory.....	34
(B) Estimation: the CES utility approach.....	37
(C) Estimation: the random utility approach.....	41
V. Unifying the Structural Hedonic and Sorting Models.....	43
(A) Identifying heterogeneous preferences from structural restrictions.....	47
(B) Identifying substitution patterns.....	52
(C) Identifying unobserved public goods and rationalizing choices.....	53

(B)	Modifying the Gibbs sampler to “jump” between disconnected sets...	103
(C)	Finding starting values in each preference set.....	104
VI.	Conclusions.....	106
4.	CHOICE SET AND DATA.....	117
I.	Introduction.....	118
II.	Generating the Data: An Overview.....	118
III.	Defining the Choice Set.....	120
(A)	Using evidence on mobility to define the external boundaries of the study region.....	121
(B)	Subdividing the study region into a set of job-house locations.....	123
IV.	Defining the Attributes that Differentiate Location Choices.....	126
(A)	Individual housing transactions.....	126
(B)	Air quality: ozone.....	129
(C)	School quality: academic performance index.....	130
(D)	Unobserved public goods.....	131
(E)	Commute time.....	132
(F)	Wages.....	133
V.	Observing Location Choices by Income and Occupation.....	135
VI.	Conclusions.....	140
5.	ESTIMATION RESULTS.....	157
I.	Introduction.....	158
II.	Overview of the Estimation.....	159

III.	Estimation: First-Stage Results.....	161
	(A) The price of housing (p_1, \dots, p_J).....	161
	(B) Housing demand parameters (β, η, ν).....	163
IV.	Estimation: Second-Stage Results for the Vertical (Epple-Sieg) Model.....	166
	(A) Overview of the GMM approach to estimation.....	166
	(B) Results from GMM estimation.....	168
V.	Estimation: Second-Stage Results for Horizontal Models.....	171
	(A) Overview of the second stage estimation strategy.....	172
	(B) CES parameter (ρ).....	175
	(C) Composite unobserved public good ($\bar{\xi}_1, \dots, \bar{\xi}_J$).....	175
	(D) Partitioning preference space $A(\alpha, \gamma, \theta)$	179
	i. single-market results.....	181
	ii. dual-market results.....	184
VI.	Quantifying the Economic Implications of the Model.....	186
	(A) Generating a distribution of preferences for the population of households.....	187
	(B) Marginal willingness-to-pay for an air quality improvement.....	189
VII.	Conclusions.....	194
6.	GENERAL EQUILIBRIUM POLICY ANALYSIS.....	216
I.	Introduction.....	217
II.	General Equilibrium Concepts and Welfare Measures.....	218
	(A) Dual-market locational equilibrium.....	218
	(B) Versions of the dual-market locational equilibrium used for policy analysis.....	220
III.	Developing a Framework for General Equilibrium Simulation.....	222
	(A) Defining the components of the general equilibrium framework.....	223

(B)	Solving for a new equilibrium: single-market vertical model.....	225
(C)	Solving for a new equilibrium: single-market horizontal model.....	227
(D)	Solving for a new equilibrium: dual-market horizontal model.....	230
(E)	Implementation issues.....	231
IV.	Using an Example to Investigate Implementation Issues.....	233
(A)	Computational burden.....	233
i.	using a subset of the population.....	234
ii.	using a subset of communities.....	236
(B)	Multiple equilibria.....	237
(C)	Summary and caveats.....	239
V.	Simulating the Response to New 8-hour Ozone Standards.....	241
(A)	Policy background.....	241
(B)	Developing the attainment scenario.....	243
(C)	Overview of the simulation.....	245
(D)	Simulation results: single-market.....	246
(E)	Simulation results: dual-market.....	250
VI.	Conclusions.....	252
7.	KEY FINDINGS AND FUTURE RESEARCH.....	271
I.	Summary of Key Findings.....	272
(A)	Conceptual and methodological results.....	273
(B)	Econometric results.....	274
(C)	Empirical results.....	275
(D)	Policy results.....	277
II.	Future Research on the Location Choice Problem.....	278
(A)	Reinterpreting the index of unobserved public goods ($\bar{\xi}_j$).....	279
(B)	Using Census micro data to describe households and their choices.....	281
(C)	Bounding the partition of preference space.....	283
(D)	Moving costs and expanding the geographic scope of the	

study region.....	286
(E) Preference heterogeneity, capitalization, and reduced-form estimation.....	288
III. Broader Implications for Future Research.....	289
(A) Structural discrete-choice models / recreation demand.....	290
(B) Multidimensional equilibria.....	291
(C) Heterogeneity and sorting by firms.....	292
IV. Conclusions.....	293
8. REFERENCES.....	298
I. Journal Articles, Books, and Working Papers.....	299
II. Data Sources.....	306
9. APPENDICES.....	308
(A) Proof of theorem 1.....	309
(B) Genetic algorithm.....	313

LIST OF TABLES

	Page
CHAPTER 1	
1.1 Factors influencing location choices for U.S. households, 2001.....	8
CHAPTER 2	
2.1 Summary of notation.....	64
2.2 Restrictions used to recover the demand for a public good from a hedonic model.....	65
2.3 An example of housing prices and air quality in four communities.....	66
2.4 An example of prices and two public goods in four communities.....	66
2.5 Adding a fifth community to the choice set.....	66
2.6 Restrictions used to recover preferences in structural hedonic and sorting models.....	67
CHAPTER 3	
3.1 Comparison of dual-market estimator to alternative sorting models.....	108
CHAPTER 4	
4.1 Land use, employment, and migration in Northern California MSAs, 1995-2000.....	143
4.2 Work location, by home PMSA, 2000.....	144
4.3 Summary statistics for housing sales, 1995-2005 (N=540,624).....	145
4.4 Variables used to convert housing prices into an annualized rental price.....	145
4.5 Mean annual wages (\$1000) by occupational category, 1999-2001.....	146
4.6 Mean annual wage for selected occupations by PMSA, 2001-2002.....	147
4.7 Number of single, dual, and multi-earner households, by PMSA.....	148
4.8 An example of the required data.....	148
4.9 Data and assumptions used in the data-generating process.....	149
4.10 Combining data on commuting patterns by occupation with data on wage income, an example.....	150
4.11 An example of the conditional income distributions reported in the CTPP special tabulation.....	150
4.12A Generating income distributions by occupation and location.....	151
4.12B Converting the universe of <i>workers in households</i> to <i>working households</i>	151
4.13 Sample of the data used in estimation: sales & related occupations	

	(SOC 410).....	152
4.14	Descriptive statistics for 122 housing communities.....	153

CHAPTER 5

5.1	Hedonic estimation of housing price indices.....	196
5.2	Housing price index by primary metropolitan statistical area.....	196
5.3	Housing demand parameters (standard errors in parenthesis).....	197
5.4	Results from GMM estimation of single-market vertical model.....	198
5.5	Hayward Unified is dominated by Vallejo City, Windsor, and Benicia.....	198
5.6	Composite unobserved public good, summarized by PMSA.....	199
5.7	Adding $\bar{\xi}$ explains the choice to live in Hayward.....	199
5.8	Summary statistics for heterogeneous preference parameters.....	200
5.9	Identifying the MWTP for improved air quality, horizontal models.....	201
5.10	MWTP for improved air quality, 3 sorting models.....	201
5.11	Average per/household MWTP, by occupation and work PMSA.....	202

CHAPTER 6

6.1	Sensitivity of simulation results to using a fraction of the sample size.....	255
6.2	Sensitivity of the simulation results to using a subset of communities.....	256
6.3A	Sensitivity of equilibria to the order of adjustment in the HEP-HEW algorithm.....	257
6.3B	Min, max, and mean changes in wages, by PMSA (for all 22 occupational categories).....	257
6.3C	Comparing the welfare of working and retired households.....	258
6.4	Summary of the attainment scenario, by PMSA.....	259
6.5A	Summary of average welfare effects and price changes, by PMSA (single-market horizontal).....	260
6.5B	Average welfare effects and price changes for individual communities (single-market horizontal).....	260
6.6A	Summary of average welfare effects and price changes, by PMSA (single-market vertical).....	261
6.6B	Average welfare effects and price changes for individual communities (single-market vertical).....	261
6.7A	Average welfare effects and price changes, by PMSA (dual-market, free mobility).....	262

6.7B	Average changes in salary and welfare by occupation and job location (dual-market free mobility).....	263
6.8	Average welfare effects and price changes, by PMSA (dual-market, limited mobility).....	264
6.9	Average per/household change in welfare, rents, and wages by equilibrium concept.....	265
CHAPTER 7		
7.1	Four community example revisited.....	295

LIST OF FIGURES

	Page
 CHAPTER 2	
2.1 A stylized depiction of the urban landscape: 4 metropolitan areas.....	68
2.2 Defining locations, markets, and mobility in hedonic and sorting models.....	69
2.3 Bid functions for housing as a function of x_1 in hedonic equilibrium.....	70
2.4 Implicit price function for x_1 and demand curves for two households.....	70
2.5 The single-crossing condition: indifference curves for three households.....	71
2.6 Partition of households across communities by income.....	71
2.7 Partition of households into communities in Epple-Sieg (1999).....	72
2.8 Revealed preference logic in characteristics space.....	73
2.9 Partition of α at $y=\$50,000$	73
2.10A Boundary indifference loci and community ranking by utility.....	74
2.10B Partitioning preference space: horizontally differentiated preferences.....	74
2.11 Metropolitan statistical areas in the United States.....	75
 CHAPTER 3	
3.1 Three communities in geographic space, attribute space, and preference space.....	109
3.2 Four communities in geographic space, attribute space, and preference space.....	110
3.3 Four communities in preference space with and without a \$350 cost of moving to community D.....	111
3.4 Aggregating communities A and B from figure 3.2.....	112
3.5 Overview of the econometric model.....	113
3.6 Using a kernel estimator to recover the unobserved public good $(\bar{\xi})$	114
3.7 Gibbs sampling algorithm.....	115
3.8 Convergence problems with the Gibbs sampler in region A_z of the partition.....	115
3.9 $A_j(\alpha, \gamma)$ is disconnected	116
 CHAPTER 4	
4.1 California's metropolitan statistical areas.....	154
4.2 The regional landscape.....	155
4.3A Ozone by school district (parts per million).....	156

4.3B	Academic performance index by school district.....	156
CHAPTER 5		
5.1	Overview of the estimation.....	203
5.2	Distribution of housing prices.....	204
5.3	Implied partition of households across communities, given $y = \$50,000$	205
5.4	Recovering an approximation to the partition of preference space.....	206
5.5	The objective function used to solve for ρ	207
5.6A	Stratification of households into ordered communities, horizontal model.....	208
5.6B	Preference regions for households with income=\$55,000.....	208
5.7A	Communities in attribute space.....	209
5.7B	Preference regions for 3 communities, single-market horizontal model.....	209
5.8	Relative preferences for air and school quality in the single and dual-market versions of the model.....	210
5.9	Stratification by relative preferences with and without job opportunities.....	211
5.10	Bounds on the marginal distribution of each heterogeneous preference parameter.....	212
5.11	Comparison between preference distributions in the single and dual-market versions of the model.....	213
5.12	Air quality and average MWTP per community.....	214
5.13	Air quality and average income, by community.....	215
5.14	MWTP for an air quality improvement by income.....	215
CHAPTER 6		
6.1	A divergent general equilibrium solution.....	266
6.2	Distribution of labor market mobility.....	267
6.3	Ozone concentrations in San Francisco and Sacramento, 1980-2003.....	268
6.4A	Ozone (ppm) in the San Francisco-Sacramento region, 1999-2001.....	269
6.4B	The study region in attainment of the 8-hour ozone standard (0.07 ppm).....	269
6.5	Average WTPGE across communities, by air quality, prices, and depiction of preference heterogeneity.....	270
CHAPTER 7		
7.1	Using “outside communities” to bound the partition of preference space.....	296
7.2	Partition of preference space with and without moving costs.....	297

Chapter 1: Introduction

Markets thrive because people are different. Access to differentiated products allows heterogeneous consumers to satisfy their diverse tastes. Over the past decade, microeconomic models have sought to use the logic of revealed preferences to exploit the ways in which the combined differences in people and products influence market outcomes. These models start by specifying the relationship between consumers and their choices in the marketplace. Then, using the observable attributes of people and their selected commodities, they invert this relationship to infer tastes for product characteristics (e.g. Berry, Levinsohn and Pakes [1995], Bajari and Benkard [2005]). Methodological advances in this process have created opportunities to advance our understanding of the demand for local public goods.

Sorting models exemplify the recent advances. They are the first revealed preference framework capable of simultaneously measuring preferences for spatially differentiated public goods at the level of an individual household and consistently analyzing the welfare effects from large scale changes in those public goods. These models demonstrate how the properties of equilibrium in the housing market can help to identify and estimate households' preferences for the local public goods that differentiate communities. Epple and Sieg (1999) offered the first illustration of this logic. In their analysis, households choose where to live based on their (exogenous) income and their preferences for the unique bundle of local public goods provided by each of a discrete set of urban communities. While households are depicted as differing in their tastes for the bundle of public goods, they all evaluate its constituent elements in the same way. This feature, labeled vertical differentiation, implies all households agree on a single ranking of communities by an index of the public goods they provide.

Relaxing vertical differentiation is important because it is reasonable to expect that different households will evaluate components of a vector of local public goods quite differently. For example, households with school age children may be more concerned about school quality while retirees may place more emphasis on climate and other environmental amenities. While several microeconomic strategies have been proposed for the situation where households differ in their relative preferences (i.e. horizontal differentiation), none have used the properties of a market equilibrium to recover preferences in a way that is

consistent with equilibrium capitalization of local public goods (Starrett [1981] and Scotchmer [1985]).

Equally important is the need to recognize that working households make two related location choices—the choice of a house and the choice of a job. Rosen (1979) suggested that because households can make adjustments in both markets, we should expect both wage rates and house prices to reflect the demand for local public goods. Despite empirical evidence in support of Rosen’s insight, most economists have focused exclusively on the housing component of location choice as a means to infer households’ valuation of amenities. The few existing studies that consider adjustment in both markets use reduced form models that restrict preferences to be homogeneous and limit the analysis to marginal changes in public goods.

By ignoring the possibilities for labor market adjustment, existing sorting models may systematically bias estimates of the economic value for changes in local public goods. The importance of this omission is underscored by the observation that proximity to employment is the reason most frequently cited by households for choosing to live in their current neighborhood, as reported in the American Housing Survey (table 1.1). While local public goods also appear to be important, they are less frequently cited as the primary factor determining households’ location choices. Similarly, in a comprehensive study of the composition of U.S. counties between 1850 and 1990, Rhode and Strumpf (2003) conclude that local public goods influence where households choose to live, but are not the dominant factor.

This research extends the existing sorting literature by developing a “dual-market” framework that uses information on households and their location choices in the housing and labor markets to infer their demand for local public goods. It also recognizes that households may have horizontally differentiated preferences for those public goods. Four interrelated objectives are addressed. First, the analysis develops a theoretical model that relates households’ location choices in the housing and labor markets to the levels of local public goods. Second, it generalizes Epple and Sieg’s (1999) empirical model to recognize that: (i) households make a joint job-house choice and (ii) households differ in their job skills and in their relative preferences for different public goods. Third, the analysis uses the empirical

model to develop a general equilibrium framework to simulate how households and markets would adjust to a large-scale change in the provision of public goods. Finally, the new dual-market estimation and simulation frameworks are used to analyze the welfare implications of a hypothetical air quality improvement that would allow the San Francisco-Sacramento region of Northern California to meet the state's recently revised standards for ambient concentrations of ground-level ozone. The following chapters describe how these four objectives are addressed. They also place this research within the context of the existing literature.

Chapter 2 reviews the literature on revealed preference models of location choice. These models start from the maintained assumption that households choose where to live based, in part, on the provision of local public goods. This premise underlies hedonic and sorting models that depict a "locational equilibrium" where each household occupies its utility-maximizing location, and all housing markets clear. Both models use the properties of this equilibrium to infer households' preferences for local public goods. The various econometric strategies for estimating each class of model are described using a unified notational framework, with special attention to the restrictions they place on the structure of preferences and on the process generating the data. After this summary, the properties of the two models are used to argue that, if the analyst chooses to specify the utility function, the resulting structural estimator could accurately be labeled a "hedonic" model or a "sorting" model. In other words, structural versions of the two models are isomorphic. Next, the location choice problem underlying these models is extended to recognize that working households choose both job and house locations. Unlike the reduced-form hedonic models that are consistent with a joint job-house choice (e.g. Roback [1982]), the characterization of the spatial landscape recognizes that workers may choose to commute between metropolitan areas. This provides a new source of information that can be used to help identify preferences.

Assuming households choose their utility-maximizing locations and all housing and labor markets clear, the properties of this "dual-market locational equilibrium" can be used to infer households' preferences for local public goods. Chapter 3 outlines an empirical framework using the logic of revealed preferences to translate information on households and

their location choices into structural parameters that represent their idiosyncratic preferences and job skills. This process starts by generalizing Epple and Sieg's (1999) parameterization for the indirect utility function to depict working households with horizontally differentiated preferences and job skills making a joint job-house choice. Given the specification for utility, the definition for the set of choice alternatives will determine what observed location choices reveal about preferences. The underlying theoretical model implies some guidelines for how this choice set is defined. After discussing these guidelines, a two-stage econometric model is developed that uses all the available data on households and their location choices to recover the parameters of the indirect utility function. A key feature of the new econometric model is that it relaxes the need for a priori assumptions on the shape of the distributions used to characterize sources for unobserved heterogeneity. This feature enables the analyst to test the sensitivity of welfare measures to (arbitrary) distributional assumptions.

The empirical model is estimated using data on the 3.2 million households living in Northern California's two largest population centers: the San Francisco and Sacramento Consolidated Metropolitan Statistical Areas. For a relatively small geographic area, this region has tremendous diversity in housing prices, provision of public goods, and job opportunities. In particular, air quality varies widely within the region. Some of the coastal areas near San Francisco have very clean air due to natural weather patterns that limit the accumulation of ground level ozone, whereas the Sacramento area is one of four most polluted in the nation. Chapter 4 describes the data that were used to define the set of location choices and characterize the population of households. The region is divided into 122 housing communities and 8 work destinations, and each (community, worksite) pair is assigned a price of housing, a set of public goods, a set of wage rates, and a commute time. The empirical model explains the location choices made by households in each of 22 occupational categories, where wage options differ across each category and non-wage income differs across workers within each category.

The fifth chapter presents results from the estimation process. To evaluate the economic implications of generalizing Epple and Sieg's empirical model, the results from the new dual-market estimator are compared to the results from two special cases—the Epple-

Sieg model, and an intermediate version of the model that admits horizontal differentiation but treats wage income as exogenous. The three models are compared in terms of their implications for households' marginal willingness-to-pay (MWTP) for a small improvement in air quality. Under baseline distributional assumptions, moving from the Epple-Sieg model to the new dual-market estimator increases the average MWTP in the population of households by approximately 47%. This estimate is an aggregation over much larger differences for households living in particular communities. Moreover, under alternative "extreme" distributional assumptions, the difference in the average MWTP between the two models can be much larger in absolute magnitude, ranging from -60% to 172%.

The estimation results can be used to generate a distribution that describes households' *ex ante* willingness-to-pay for a large-scale air quality improvement in the San Francisco-Sacramento region. However, a sufficiently large improvement will induce some households to move to a new house and perhaps to a new job, placing upward pressure on housing prices and downward pressure on wage rates in the improved locations. Subsequent adjustments in prices and wages needed to clear those markets will have additional welfare implications. Chapter 6 develops a framework to simulate this process of "general equilibrium" adjustment. The prices, wages, and location choices that define the new dual-market locational equilibrium can then be used to generate a distribution that describes households' *ex post* willingness-to-pay for the improvement. As in chapter 5, the economic implications of moving to a dual-market framework are evaluated by comparing the simulation results from the dual-market model to the simulation results from selected special cases that restrict the depiction of preference heterogeneity and the opportunities for labor market adjustment. Each version of the model is used to analyze the welfare implications for a hypothetical reduction in emissions of ozone precursors that would allow the entire San Francisco-Sacramento region to meet the state's new limit on ambient ozone concentrations of 0.07 parts per million, set in May, 2006.

Finally, chapter 7 summarizes the key findings from this dissertation and suggests directions for future research. While the idea of using households' location choices to reveal their demand for spatially differentiated public goods dates back at least to Ridker's (1967) air pollution study, structural sorting models appear to be at an early stage of development.

The results from this dissertation identify areas where additional advances in microeconomic methods are needed to improve the way location choices can be used to infer the demand for local public goods. In addition, some of the results offer opportunities to extend this research beyond the location choice problem. The chapter concludes by discussing the broader implications for future research in public and environmental economics.

Table 1.1: Factors Influencing Location Choices for U.S. Households, 2001

Main Reason for Choice of Present Neighborhood	Percentage
Convenient to job	21.3%
Other	17.9%
House was most important consideration	14.8%
Convenient to friends or relatives	14.2%
Looks/design of neighborhood	14.0%
Good schools	6.6%
All reported reasons equal	3.4%
Not reported	3.0%
Convenient to leisure activities	2.3%
Other public services	1.2%
Convenient to public transportation	1.2%

Source: U.S. Census Bureau, American Housing Survey Branch, American Housing Survey, 2001. Table 2-11.

Chapter 2: Revealed Preference Models of Location Choice

I. Introduction

This chapter reviews the literature on revealed preference models of location choice. These models start from the maintained assumption that households choose where to live based, in part, on the provision of local public goods. A “locational equilibrium” arises when the housing market clears and no household can improve their location given their available resources. Empirical models use the properties of a locational equilibrium to relate consumers’ observable choices to their unobserved demand for public goods. Finally, the process yields a consistent econometric framework that provides estimates of the demand for location specific public goods. This chapter describes the theoretical, empirical, and econometric models of location choice that allow an analyst to recover an individual household’s demand for a public good.

The literature is divided into two approaches: hedonic models and sorting models. While their underlying theories differ, they apply a common strategy. They assume households “pay” for local public goods through the price of housing and then use data on housing purchases and spatial variation in public goods to infer the demand. This strategy presents a fundamental identification problem. Since each household typically makes a single housing purchase it is possible to identify, at most, one point on that household’s demand curve. In order to recover the entire demand curve, the econometrician must provide some additional information about the structure of preferences. Recent advances in microeconomic methods have created a great deal of flexibility in how this information can be provided. Hedonic and sorting models differ in how they have exploited these advances.

The first three sections of the chapter review the existing literature. Section II describes how models of location choice characterize the urban landscape and formalizes a household’s location choice problem. This provides a common foundation for the hedonic and sorting models in the following two sections. Section III begins with a brief overview of hedonic theory before summarizing the different econometric strategies that have been developed. Section IV provides a parallel summary of the sorting literature. While the suite of models discussed in sections III and IV all start from the same choice problem, they differ

in three respects: (a) how they define what constitutes a “location”, (b) how they specify the geographic scope of the choice set, and (c) in the restrictions they impose to identify preferences.

The last two sections of the chapter seek to unify parts of the existing literature and suggest an extension to the underlying location choice problem. Section V begins by suggesting that structural hedonic and sorting estimators are, in principle, isomorphic. The similarity between them is demonstrated by decomposing the revealed preference logic that drives the estimators and comparing the restrictions they impose on preferences. Finally, section VI redefines the location choice problem to link preferences for public goods to location choices in both the housing and labor markets. Implementing this logic in a structural framework requires a new depiction of the urban landscape. An advantage of this new perspective is that it can exploit more sources of information to identify preferences. However, it also introduces a new set of econometric and data challenges, which are addressed in chapters 3 and 4.

II. Choice and Local Public Goods

All the models in this chapter start from the same basic choice problem: the availability of housing and public goods varies across an urban landscape and each household chooses to occupy the location in that landscape that provides its preferred bundle of goods, given the wage it can earn, its non-wage income, and the relative prices involved. Every household pays for its location choice through the price of housing. Working households may also pay indirectly through the wages they earn. In order to link a household’s location choice to its demand for an individual public good, the problem is formalized using the characteristics approach to consumer theory (Lancaster [1966], Gorman [1980]). That is, the utility a household obtains from each location is written as a function of the characteristics of that location. This keeps the dimensionality of the problem manageable and allows locations to be treated as a differentiated product. Table 2.1 provides a quick reference to the notation that will be used.

Let the urban landscape consist of $j = 1, \dots, J$ locations, each of which is defined by a bundle of housing characteristics and public goods. Let h_j be a vector of structural characteristics that fully describe housing at location j . h_j includes the number of bedrooms, the number of bathrooms, square feet, and lot size, to name a few. Let g_j denote a vector of local public goods which are consumed exclusively by the occupants of location j (e.g. crime, school quality, air quality, open space). “Public goods” are defined here to include environmental amenities in addition to services that are produced from local taxes. Throughout this chapter the supply of public goods is assumed to be exogenous.

The landscape at this stage of the model simply consists of j locations, each of which is defined by a (h_j, g_j) combination. In a slight abuse of notation, the spatial resolution of a location varies across the models in this chapter. The j subscript will indicate a house, a community, or a metropolitan area depending on the model under discussion. While the definition of a location differs from model to model, the underlying topography of the spatial landscape remains the same. Let $k = 1, \dots, K$ index a set of metropolitan areas that characterize this landscape. Figure 2.1 shows a stylized example with four metro areas. It will be used throughout this chapter to illustrate how the different models depict an individual location and the set of possible location choices.

A household’s utility depends on the characteristics of housing and local public goods at its location and on its consumption of a composite numeraire private good, b . Households are heterogeneous. They differ in their preferences (α) and in their demographic characteristics (d) such as income, age, education, and race. Let the population of households be indexed from $i = 1, \dots, I$. Then the utility obtained by household i at location j can be represented as: $U(b, h_j, g_j; \alpha_i)$. Though a household may contain many members with different demographic characteristics and preferences, it will be treated throughout this analysis as an indivisible economic agent.

Each household is assumed to choose a location and a quantity of b that maximize its utility subject to a budget constraint, as follows:

$$(2.1) \quad \max_{j,b} U(b, h_j, g_j; \alpha_i) \text{ subject to } y_{i,j} = \hat{y}_i + w_{i,j} = b + P_j.$$

In the budget constraint, the price of the numeraire is normalized to one and P_j represents annualized expenditures on housing at location j ; in other words, P_j is the cost of occupying a single home for one year. $y_{i,j}$ is the household's total annual income. It depends on \hat{y}_i , the household's exogenous non-wage income and $w_{i,j}$, the wage income the household earns if it lives at location j . Wages are allowed to vary across the urban landscape so that, in general, income is endogenous to location choice. However, most of the models described in this chapter treat income as exogenous.

The choice problem outlined in (2.1) provides the basis for most of the models discussed here. This logic requires that each household be observed living in its utility-maximizing location. To guarantee this is the case it is useful to remove some of the potential sources of "friction" from the problem. First, all households are assumed to share the same objective evaluation of local public goods and housing characteristics. Second, households are assumed to be freely mobile within the geographic region defined as the choice set; i.e. moving costs are zero. Third, every household is assumed to face the same schedule of housing prices; i.e. there is no price discrimination.

III. Hedonic Models

Hedonic models express the price of a differentiated product as a function of its characteristics. While the conceptual basis for this approach dates back to the 1920s or 1930s, the method was popularized by Griliches's (1961) work on using a hedonic price function to make quality adjustments to price indices for automobiles. Rosen (1974) strengthened the economic foundations of the method by demonstrating that the hedonic price function can be interpreted as an equilibrium relationship resulting from the interactions between all the buyers and sellers in a market. He also suggested that the hedonic price function could be used to infer the demand for product characteristics, inspiring a line of research that continues today.

In the public, environmental and urban economics literatures, hedonic models are frequently used to derive implicit prices for local public goods. Consider air quality. While a household does not pay anyone directly for the air its members breathe, the household purchases the right to have access to this air on a regular basis through the price of its home. Regressing housing prices on air quality, while controlling for other housing and location-specific characteristics, was thought to offer the potential for revealing the marginal price that households implicitly pay for improvements in air quality. The possibility of using this information to infer the demand for air quality (and other public goods) has intrigued economists since Rosen's (1974) paper.

To formalize the discussion, the annualized price of housing can be expressed as a function of its structural characteristics and local public goods: $P = P(h, g)$. This price function can be estimated econometrically using data on housing prices and characteristics from individual real estate transactions. However, it is virtually impossible for the econometrician to observe every relevant housing characteristic and local public good. To reflect this, let $x \subset [g, h]$ represent the characteristics observed by both households and the econometrician, and $\xi \subset [g, h]$ represent characteristics that are observed by households but not by the econometrician such that: $x \cup \xi = g \cup h$. Since hedonic models treat local public goods the same as structural characteristics of housing, there is no loss of generality in using x to represent the observable dimensions of both. Using the new notation, the hedonic price function can be rewritten as the second term in (2.2).

$$(2.2) \quad P(h, g) = P(x, \xi) = P(x, B, e(\xi)).$$

The third term is an econometric approximation characterized by the parameter vector B and the residual $e(\xi)$. In order to focus the discussion on unobserved characteristics, other sources of error such as functional form misspecification and error in measuring x are assumed to be negligible. Finally, to keep the notation manageable it is helpful to concentrate on a single public good: $g_1 = x_1 \in x$. This simplification does not affect any of the conclusions.

Partially differentiating the econometric approximation to the hedonic price function

with respect to x_1 provides an estimate of the marginal price function for x_1 .

$$(2.3) \quad \hat{P}_1 = \frac{\partial P(x, B, e(\xi))}{\partial x_1}.$$

\hat{P}_1 is the marginal contribution of x_1 to the price of housing given the current level of x_1 and levels of the other characteristics. In other words, (2.3) is an implicit price function that describes the marginal price for each unit of x_1 relevant for every household in the market described by $P(h, g)$.

Implementing Rosen's proposal to estimate the demand for x_1 requires the econometrician to place some restrictions on preferences as well as on the process generating the data. Three approaches are described. Each differs in the balance they choose between structural restrictions and econometric restrictions. While the three methods differ in the information they supply, they all make two key assumptions about the choice set. First, they assume households are free to choose any combination of locational characteristics; i.e. the choice set is continuous. Second, they assume wage income is exogenous to location choice.

(A) Property Value Model: Continuous Choice Set

Hedonic property value models typically define a "location" as an individual house and assume that households are free to choose any location within the "housing market". While the concept of a "housing market" lacks an obvious spatial analog, empirical work typically uses data from a single metropolitan area or a subset thereof. However, the rule of thumb is to combine areas that a reasonable number of households would consider as alternative choices (Palmquist [2003]). With this in mind, let the four metro areas in figure 2.1 comprise 3 housing markets. That is, suppose many households consider metro areas 2 and 3 to be close substitutes. Panel A in figure 2.2 illustrates how the spatial landscape would be depicted in a hedonic property value model, letting $q = 1, \dots, Q$ index housing markets¹.

¹In general, the analyst's definition for a "housing market" plays an important role in revealed preference models of location choice. The next two chapters discuss this in more detail. Chapter 3 demonstrates how estimates for the marginal willingness-to-pay for a public good can vary with the definition chosen by the

Another distinguishing feature of hedonic property value models is that they treat wage income as exogenous and assume that households face a continuum of choices—they can choose a house with any combination of structural characteristics and public goods. Thus, the choice variable in the utility maximization problem can be changed from j to (h, g) and equation (2.1) can be rewritten as (2.4).

$$(2.4) \quad \max_{h,g,b} U(b, h, g; \alpha_i) \quad \text{subject to} \quad y_i = b + P(h, g).$$

Each household chooses the combination of structural characteristics, local public goods, and the numeraire composite commodity (b) that maximize its utility, given its budget constraint. The budget constraint differs from the one in (2.1) in two ways. First, income is fully exogenous, simplifying the expression. Second, expenditures on housing are now represented by the hedonic price function.

The key result of the hedonic property value model lies in the first order conditions to the utility maximization problem. Equation (2.5) shows the first order condition for the public good².

$$(2.5) \quad FOC : \frac{\partial P}{\partial x_1} = \frac{\partial U / \partial x_1}{\partial U / \partial b}.$$

The equation implies that consumers will maximize their utility by choosing a house that provides them with a level for the public good at which their marginal willingness-to-pay for an additional unit exactly equals its marginal implicit price. This relationship obviously requires continuity in the choice set. Figure 2.3 illustrates the first-order condition. It shows bid functions for housing in the x_1 dimension for two households. The bid functions express each household's willingness-to-pay for housing as a function of x_1 , given levels of all the other housing characteristics and public goods, and the household's preferences and income

analyst. Chapter 4 discusses how empirical data on migration and commuting patterns can be used to help define a housing market.

² Analogous first order conditions can be derived for each structural characteristic in the hedonic price function. That is, there is nothing special about the way public goods enter the hedonic model; they are treated the same as structural characteristics.

$(x_2, x_3, \dots, \xi, \alpha, y)$. Each household will select the quantity of x_1 where its bid function is tangent to the hedonic price function. In the figure, the two households purchase homes that are identical except in their provision of the public good. Household 1 spends $\$1$ on a house that provides $x_{1,1}$ units of the public good and household 2 spends $\$2$ on a house with $x_{1,2}$. Therefore $(\$2 - \$1)$ is the effective cost of consuming $(x_{1,2} - x_{1,1})$ additional units of x_1 , given the levels of all the other housing characteristics and public goods.

Rosen (1974) suggested a two-step procedure that would use the information in the first order condition to estimate the demand for a product characteristic—in this case x_1 . The first step is to use micro data on housing transactions to estimate the reduced form housing price function in (2.2) and partially differentiate it to recover \hat{P}_1 , the marginal price function for x_1 . The first order condition in (2.5) implies that by evaluating \hat{P}_1 using each household's chosen level of x_1 it is possible to recover an estimate of the marginal willingness-to-pay (MWTP) for x_1 . Combining this information with the level of x_1 at a household's location gives exactly one point on that household's demand curve. Figure 2.4 illustrates the idea. Household 1's demand curve (D_1) intersects \hat{P}_1 at the point where its MWTP exactly equals the marginal price for an extra unit of x_1 . While the marginal willingness-to-pay is identified, demand curves are not. For example, the figure shows two different demand curves that are both consistent with household 2's observed choice. An infinite number of demand curves could pass through the point observed for households 1 and 2, and the same is true for every other household. Clearly, more information is needed to identify household level demand curves.

For the second step of his procedure, Rosen suggested regressing estimates for the MWTP on product characteristics and a set of exogenous demand shifters such as income and demographic characteristics. Equation (2.6) illustrates the idea, where Ω is a parameter vector, d is a vector of observable demographic characteristics including income, and u is the residual.

$$(2.6) \quad \hat{P}_1 = f(x, d, \Omega, u).$$

The logic is that if households' unobserved preferences for x_1 are highly correlated with their demographics, the regression in (2.6) would recover the inverse demand for x_1 . It was later recognized that this logic makes two important assumptions.

First, identifying the inverse demand curve with data from a single housing market requires that there be some nonlinearity in the marginal implicit price function. Brown and Rosen (1982) demonstrated this for a case where the marginal implicit price function and the inverse demand curve are both linear. Their result is shown in (2.7), which is simply a linear representation of (2.6).

$$(2.7) \quad B_0 + B_1x = \Omega_0 + \Omega_1x + \Omega_2d + u.$$

The regression will simply recover B_1 , the parameter vector that characterizes the implicit marginal price function; there is no new information to identify the shape of the demand curve. The second qualifying assumption arises because households choose prices and quantities simultaneously. As a result x_1 will be endogenous in (2.7) (Epple [1987]; Bartik [1987]). Thus, to avoid biased estimates, instruments are required.

Overall then, to recover the demand for a public good, the econometrician must provide two additional sources of information. First, preferences must be restricted in a way that makes it possible to identify the demand curve. Second, the data generating process must be restricted in a way that makes it possible to estimate the demand curve consistently. Three general econometric strategies have been developed to address these issues.

(i) Strategy #1: Multiple Markets

The first strategy to recover the demand for x_1 implements Rosen's two-step approach by estimating hedonic price functions for multiple housing markets (Brown and Rosen ([1982])). The identifying restriction on preferences is that they are highly correlated with income and demographic characteristics. To develop this idea formally, suppose that preferences for structural characteristics of housing and local public goods can be written as a constant function of observable demographic characteristics and unobserved household-specific tastes

(ε) as in equation (2.8).

$$(2.8) \quad \alpha = f(d, \varepsilon)$$

Thus, two households with the same demographic characteristics and idiosyncratic tastes will have the same value for α_i in the objective function defined in equation (2.4). The identifying restriction is that (2.8) is separable in d and ε , with the demographic sub-function constant across each of the Q markets as shown in (2.9).

$$(2.9) \quad \alpha = f(d) + \varepsilon, \quad \text{with } f_1(d) = f_2(d) = \dots = f_Q(d).$$

Intuitively, this restriction makes it possible to obtain multiple observations on the demand curve of each household “type”. For example, in a representative application by Palmquist and Israngkura (1999) one particular household type is 40 year old, white, married couples with two children. They are restricted to have the same demand for air quality (up to ε) whether they live in San Francisco, Denver, Chicago, or 10 other housing markets in the sample. In each market, this type of household faces a different hedonic price schedule, and therefore will choose different implicit prices and qualities of air, identifying 13 different points on the common portion of their demand curve.

The restriction in (2.9) is sufficient to identify individual demand curves, even if \hat{P}_1 is linear. The econometric strategy is to estimate the hedonic price function separately in each housing market and then pool the resulting sets of estimates for the marginal implicit prices and regress them on housing characteristics and household demographics, as shown in (2.10).

$$(2.10.a) \quad P_q = f(x_q, B_q) + e(\xi_q) \quad \text{for } q = 1, \dots, Q \text{ markets.}$$

$$(2.10.b) \quad \hat{P}_1(x, B) = f(x, z, d, \Omega) + \varepsilon.$$

In order for the equation in (2.10) to provide consistent estimates of the demand for x_1 the three econometric restrictions shown in (2.11) must be satisfied.

$$(2.11.a) \quad E[\xi | x] = 0.$$

$$(2.11.b) \quad E[z'x] \text{ is full rank.}$$

$$(2.11.c) \quad E[\varepsilon | d, z] = 0.$$

First, to estimate B consistently, the unobserved characteristics must be uncorrelated with the observed characteristics (2.11.a). Second, a set of instruments (z) for x must be introduced to address the endogeneity problem (2.11.b). Third, the residual to the inverse demand function (which is interpreted as household-specific tastes) must be uncorrelated with demographic characteristics and the instruments (2.11.c).

Empirical studies have invoked the restrictions in (2.9) and (2.11) to estimate demand functions for various public goods including air and water quality, using data from between two and thirteen markets (Taylor [2003]). A common feature of these studies is that they restrict the hedonic price function to have an additively separable error term, as in (2.10.a). This is important because it implies that unobserved housing characteristics do not influence the demand for x_1 .

The multi-market approach is data intensive. To estimate the model, one must obtain micro data on the characteristics of houses and households in multiple markets, as well as good instruments for the housing characteristics. Recently, Ekeland et al. (2004) suggested an alternative strategy that would recover the demand for a housing characteristic from data on a single market, without requiring instruments. The tradeoff is that additional structure must be placed on the utility function.

(ii) Strategy #2: Restricting Preferences

Ekeland et al (2004) argue that the marginal implicit price function is generally nonlinear and will only be linear in special cases. They demonstrate that this nonlinearity provides a way to identify the demand for x_1 if the utility function satisfies some general restrictions. Most importantly, marginal utility must be monotonically separable in characteristics, demographics, and household-specific tastes so that it can be written as (2.12).

$$(2.12) \quad \partial U / \partial x = f_1(x) + f_2(d) + f_3(\varepsilon).$$

Assuming the restriction holds, Rosen's suggested two-stage strategy can be used to estimate the demand for x_1 from data on a single housing market. The first step is to estimate the additively separable hedonic price function in (2.13a) just as in the multi-market approach.

$$(2.13.a) \quad P = f(x, B) + e .$$

$$(2.13.b) \quad \hat{P}_1(x, B) = f_1(x, \Omega) + f_2(d, \Gamma) + \varepsilon .$$

In the second step, the estimated marginal implicit prices are regressed on housing characteristics and demographics, using $E[f_1(x)|d]$ as an instrument for $f_1(x)$ in (2.13.b).

The advantage of this approach is that it eliminates the need for additional instruments and data on multiple markets, relaxing (2.9) and (2.11.b). While this is an interesting extension, the theorems that support it are only proven for the case where x , d , and ε , are each 1-dimensional. This restriction precludes using the logic in most applications where x includes a large number of housing characteristics and one or more local public goods. Furthermore, Ekeland et al. do not allow for unobserved housing characteristics, which is why ξ is missing from the residual to the hedonic price function in (2.13.a). Perhaps because of these limitations, their strategy has yet to be applied to the housing market.

In theory, the requirement that x be 1-dimensional could be relaxed if the analyst were prepared to restrict the nature of preference heterogeneity. The idea would be to restrict preferences in a way that allows the bundle of housing characteristics and public goods to be consistently represented by a 1-dimensional index of overall housing "quality". This can be done if households have identical relative preferences for every pair of elements in the bundle; i.e. they must have identical values for the weights in the quality index. However, estimating Ekeland et al.'s model under this preference restriction would also require developing an econometric method for recovering the constant "weights" in the quality

index³. It is an open question whether the theorems in Ekeland et al. can be generalized to higher dimensions in order to allow more flexible forms of preference heterogeneity and to allow unobserved housing characteristics.

(iii) Strategy #3: Fully Specifying the Utility Function

The first two strategies both take a reduced-form approach to estimating the inverse demand curve that requires preferences to be highly correlated with household demographics. This assumption can be relaxed, along with the need for data on demographics, by fully specifying the shape of the utility function. As with the first two strategies, this approach starts by estimating the hedonic price function in (2.14a). However, in this case the error term need not be separable (Bajari and Benkard [2005]).

$$(2.14.a) \quad P = f(x, B, e(\xi)).$$

$$(2.14.b) \quad \hat{P}_1(x, B, e(\xi)) = U_1(x, \xi, b; \alpha).$$

Since the shape of the utility function is known, the first-order condition in (2.14.b) can be used to solve directly for the structural preference parameters. Given preferences, the demand for x_1 can be calculated directly.

What can be learned about preferences from (2.14.b) depends on what is assumed about the structure of the utility function and preference heterogeneity. For example, in another air quality application, Chattopadhyay (1999) uses Diewert's utility function to develop the first-order condition in (2.15).

$$(2.15) \quad \hat{P}_1 = \alpha_0 + \alpha_1 \sqrt{x} + \alpha_2 d + \varepsilon.$$

He constrains households to have homogeneous preferences and estimates constant values for α . More recently, Bajari and Kahn (2005) specify a log-linear utility function and

³ This issue is revisited during the discussion of sorting models in section IV. Some sorting models exploit a similar index restriction on preferences to recover households' (heterogeneous) preferences for overall housing quality and their (homogeneous) preferences for the individual elements of the quality index, while simultaneously controlling for unobserved public goods.

estimate a semiparametric hedonic price function, yielding the expression in (2.16).

$$(2.16) \quad \hat{P}_1 = \frac{\alpha_1}{x_1} \Rightarrow \alpha_1 = x_1 \hat{P}_1.$$

They allow households to have heterogeneous preferences for housing characteristics, and use the expression to solve for values of α for each individual household. That is, substituting the household's chosen level of x_1 and the other characteristics into $\hat{P}_1(x, B, e(\xi))$ allows them to solve for α_i . Thus, as one might expect, with a more rigid structure for the utility function Bajari and Kahn are able to recover more information about preference heterogeneity.

An advantage of the structural approach is that data on demographics are not needed. Nonetheless, in some policy applications the correlation between preferences and demographics may be important. For example, one might hypothesize that preferences for school quality differ for households with and without school-age children. If demographic information on households is available, the structural preference parameters estimated from (2.16) can be regressed on demographics. For example, Bajari and Kahn (2005) estimate the additively separable function in (2.17).

$$(2.17) \quad \alpha = f(d) + \varepsilon.$$

While the flexibility of the approach used by Bajari and Kahn (2005) is appealing, it has two limitations. First, it can globally identify only as many structural parameters as there are observable product characteristics. Second, in order to recover preferences for individual households, the mean independence assumption in (2.11.a) must be strengthened. These two restrictions are presented in equations (2.18.a) and (2.18.b).

$$(2.18.a) \quad U(x, \xi, b; \alpha) \text{ is fully specified with: } \dim(\alpha) \leq \dim(x)$$

$$(2.18.b) \quad x \text{ and } \xi \text{ are independent}$$

The independence assumption can be relaxed if there are instruments available for the

endogenous characteristics.

(iv) Comparing the Three Strategies

While each of the three strategies presents a different solution to the identification problem shown in figure 2.4, they all require structural restrictions on preferences and econometric restrictions on unobserved elements of the data. These restrictions are difficult to verify. Columns 1 to 3 of table 2.2 summarize the essential restrictions for each strategy. The multi-market approach and the fully specified approach are the two extremes. The former is data intensive and relies primarily on econometric restrictions; the later only requires data from a single market and relies primarily on structural restrictions. The strategy proposed by Ekeland, Heckman, and Nesheim (2004) seems to offer a middle ground by placing general restrictions on preferences and using data from a single market. However, until their strategy can be extended to allow for unobserved characteristics and more than one observed characteristic, its applicability to recover the demand for local public goods is limited. Moreover, the requirement that marginal utility be monotonically separable in x , d , and ε , is often violated by relatively simple functional forms such as the Diewert utility function used by Chattopadhyay and the log-linear utility function used by Bajari and Kahn⁴.

In addition to the identification and endogeneity issues that motivated the three strategies, there are a number of practical econometric issues to consider in estimating the demand for x_1 . For example, unobserved characteristics may disproportionately affect the performance of different parametric specifications for the hedonic price function (Cropper et al. [1988]). In reduced-form applications, researchers often assume that the presence of unobserved characteristics causes the error terms to be spatially correlated, leading to inefficient parameter estimates⁵. In theory, this can be corrected with a weighting matrix. However, the standard weighting techniques only address very specific patterns of spatial correlation. Moreover, the researcher must know the form of spatial correlation in advance.

⁴ In Bajari and Kahn's (2005) specification, the nonseparability can be seen by combining (2.16) and (2.17).

⁵ Spatially correlated error terms in the hedonic price function could also be caused by households that systematically sort across the urban landscape according to their idiosyncratic tastes for individual locations.

There is a large literature on spatial error correlation and other econometric issues associated with obtaining efficient estimates of parameters in reduced form hedonic models⁶.

All three of the strategies share the maintained assumption that households face a continuous set of choices for local public goods and housing characteristics. While this may be a reasonable assumption for structural characteristics like square feet and lot size, it seems less appropriate for some public goods. For example, school quality varies discretely across school districts. Similarly, access to public pools, tennis courts and community centers may be limited to homeowners in a subdivision. Spatial bundling of public goods could be another source of discreteness in the choice set. In order to live within walking distance of downtown dining and other cultural amenities, a household may have to tolerate congestion, noise and light pollution in the evenings. Likewise, houses near the edge of a city may have scenic views and access to open space bundled together with disamenities from productive farming such as noise, dust, and odors. Thus, in this example, “scenic views” and “immediate access to downtown amenities” are mutually exclusive public goods. While each may be chosen continuously, selecting a high value for one limits the range of possible choices for the other.

(B) Property Value Model: Discrete Choice Set

If households are not free to choose continuous quantities of x_1 and the other locational characteristics, the first-order conditions in (2.5) no longer characterize the choice process. This discreteness undermines the traditional hedonic method. If equation (2.5) fails to hold, marginal implicit prices will not reveal households’ marginal willingness-to-pay. However, a household’s location choice will still provide information about its preferences. This information can be extracted if the utility function is fully specified.

When the choice set is discrete, utility maximization is characterized by the set of inequalities in (2.19) rather than the usual first-order conditions⁷.

⁶ Palmquist (2003) summarizes some of this work.

⁷ Income and prices enter the utility function by substituting for the numeraire private good.

$$(2.19) \quad U(h_j, g_j, y_i - P_j; \alpha_i) \geq U(h_n, g_n, y_i - P_n; \alpha_i), \quad \forall n = 1, \dots, J.$$

The equation simply says that if household i chooses location j it is because that location provides it with at least as much utility as any other location. Given a parametric form for the utility function, the inequalities serve to identify a set of parameter values that are consistent with each observed choice.

To consider how “set identification” works, consider the simple example in table 2.3 where households choose among four house locations which are differentiated by prices and air quality. Using the log-linear utility function from Bajari and Kahn (2005), it is straightforward to solve for a set of values for α that are consistent with each choice. For households in location 2, the inequalities in (2.19) can be written as (2.20).

$$(2.20) \quad \begin{aligned} \alpha \ln(air_2) - P_2 &\geq \alpha \ln(air_1) - P_1 \\ \alpha \ln(air_2) - P_2 &\geq \alpha \ln(air_3) - P_3 \\ \alpha \ln(air_2) - P_2 &\geq \alpha \ln(air_4) - P_4 \end{aligned} .$$

Inserting values from the table and solving the system reveals that households in location 2 must have values for α between 0.637 and 3.207. Any value for α within this range is consistent with the choice of community 2. Thus, the choice of community 2 “set identifies” α to be within that range.

The table also reveals a problem. Location 3 is more expensive than location 2 and it has lower air quality. This would seem to imply that households living in location 3 dislike clear air, which seems quite unlikely. A more likely explanation is that the price in location 3 reflects an important unobserved characteristic. This logic underscores the importance of accounting for unobserved characteristics in the estimation.

Bajari and Benkard (2005) develop a revised version of Rosen’s two step strategy that explicitly accounts for unobserved characteristics. Their method starts by using the hedonic price function to recover values for the unobserved characteristics, and then it uses those values to rationalize observed choices, solving the system in (2.19). Their approach requires that the specification for the utility function satisfy the three restrictions in (2.21).

$$(2.21.a) \quad U(x, \xi, b) \text{ is continuously differentiable in } b \text{ with } \frac{\partial U(x, \xi, b)}{\partial b} > 0$$

$$(2.21.b) \quad U(x, \xi, b) \text{ is Lipschitz continuous in } (x, \xi).$$

$$(2.21.c) \quad \frac{\partial U(x, \xi, b)}{\partial \xi} > 0 .$$

These three restrictions, together with the independence assumption from (2.18.b), make it possible to recover a 1-dimensional unobserved characteristic at each location. The composite unobserved characteristic, $\bar{\xi}_j = f(\xi_j)$, can be interpreted as an index of all the unobserved characteristics at each location. It is recovered from the joint distribution of prices and observed characteristics $F(x, P)$. Specifically, Bajari and Benkard prove that (2.22) must hold.

$$(2.22) \quad F_{P|x=x_j}(P_j) = \bar{\xi}_j .$$

In words, the composite unobserved characteristic in location j equals the quantile of P_j in the distribution of prices conditional on observed characteristics. Their proof relies on a monotonic transformation of $\bar{\xi}$ such that $\bar{\xi} \sim U(0,1)$. This is not restrictive; since the unobserved characteristic has no natural units, it is only defined up to a monotonic transformation.

Once values for the unobserved characteristic have been recovered, they can be inserted into the system of inequalities in (2.19) to solve for sets of values for the heterogeneous parameters that explain observed location choices. To move from these preference sets to a point value for each household, a distribution must be specified for each heterogeneous preference parameter⁸. Just as when the choice set is continuous, the demand for x_1 can be computed directly after the preference parameters have been recovered.

Column 4 in table 2.2 summarizes the restrictions required for Bajari and Benkard's

⁸ Section V discusses this process in greater detail, along with other important aspects of Bajari and Benkard (2005).

econometric strategy. To date, there have been no applications to the housing market. It is worth noting that their strategy does not require the choice set to be fully discrete. If the utility function is separable in a characteristic that can be chosen continuously, the hedonic first-order condition could be used to solve for the corresponding preference parameter. The other parameters could be recovered using the system of inequalities. That is, one could estimate a structural discrete-continuous hedonic model.

(C) *Interregional Hedonic Model*

All of the hedonic property value models treat wage income as exogenous. However, for working households, there are two related components of location choice—the choice of a house and the choice of a job. Rosen (1979) suggested that because households can make adjustments in both markets, we should expect both wage rates and house prices to reflect variation in local public goods. Intuitively, locations providing fewer public goods must offer higher wages and lower housing prices to induce working households to locate there. As more working households locate in high-amenity areas, the supply of labor and the demand for housing will increase, most likely resulting in lower wages and higher rents. Thus, public goods may influence spatial variation in wages and rents which, in turn, will affect the location choices of individual households. Roback (1982) formalized this intuition and demonstrated that spatial variation in both rents and wages should be used to calculate the implicit price of a public good.

Roback's "interregional" hedonic model differs from the property value hedonic in the way it depicts the spatial landscape. Locations are defined as metropolitan areas and the choice set is national. Each of the $j = 1, \dots, K$ metro areas is assumed to offer a different set of public goods and job opportunities, which are homogeneous within that location. While households are free to move between metro areas, the theory does not allow commuting between them. Restricting people to live and work in the same place enforces the theoretical link between prices, wages, and public goods at each location. Figure 2.2B illustrates this landscape.

Households choose to live in their utility maximizing location, subject to a budget

constraint that depends on the income they would earn in that location. Equation (2.23) illustrates this problem.

$$(2.23) \quad \max_{j, \bar{h}, b} U(b, \bar{h}, g_j; \alpha) \text{ subject to } y_{i,j} = \hat{y}_i + w_{i,j} = b + p_j \cdot \bar{h}.$$

In the budget constraint, p_j represents the annualized price of a homogeneous unit of housing in metro area j ; i.e. the rental rate. The bar superscript on \bar{h} is meant to differentiate the concept of a homogeneous unit of housing from h , which continues to represent a vector of structural characteristics. Unlike the property value hedonic, households are restricted to have homogeneous preferences: $\alpha_1 = \dots = \alpha_I$. This simplifies the problem by requiring every household to be indifferent between all possible locations in an equilibrium, in which case indirect utility must be constant across space.

$$(2.24) \quad V_i(w_j, p_j; g_j) = \bar{V} \quad \forall j = 1, \dots, K, \quad i = 1, \dots, I.$$

By totally differentiating (2.24) and applying Roy's identity, the marginal willingness-to-pay for the public good x_1 can be written as (2.25).

$$(2.25) \quad \frac{\partial V / \partial x_1}{\partial V / \partial w} = \bar{h} \frac{dp}{dx_1} - \frac{dw}{dx_1}.$$

As in the property value hedonic, every household will choose a location that sets its MWTP for x_1 equal to the implicit marginal price. The difference is that the interregional model defines the total implicit marginal price for x_1 as the sum of its implicit marginal prices in the housing and labor markets. In other words, households "pay" for the public goods in their metro area through both rents and wages.

After converting housing prices into a rental rate for each metropolitan area, the hedonic price function can be estimated just as in the property value models (Blomquist et al.

[1988]).⁹ The hedonic wage function is based on a similar logic. The wage rate each household earns is assumed to be a function of its demographic characteristics and local public goods. The wage function is shown in (2.26), where Δ is a vector of parameters characterizing the econometric approximation to the function and e is the error term.

$$(2.26) \quad w(d, g) = w(d, g, \Delta, e).$$

Partially differentiating the wage function with respect to $x_1 \in g$ yields its marginal implicit price function. Combining the two marginal implicit price functions calculated from rents and wages provides the total implicit price function for the public good, as in (2.25). Since households are assumed to be homogeneous, this implicit price function also represents an inverse demand curve for x_1 .

Empirical applications have focused on using the interregional hedonic model to calculate “quality-of-life” indices that compare the composite provision of public goods across U.S. metropolitan areas. This line of research begins by choosing parametric forms for the hedonic price and wage functions that imply a constant marginal implicit price for each public good. Then, the estimates for the implicit prices are treated as “weights” in a linear index of public goods provision. In these applications, both housing prices and wages have been found to reflect a substantial share of the total implicit price of many public goods (e.g. Blomquist et al. [1988]; Graves and Waldman [1991]).

The maintained assumption that households share an identical demand for public goods is recognized as a major limitation of the interregional hedonic model. Roback (1988) relaxes this restriction by generalizing her model to depict households with two different job skills which are complements in production. While households are still assumed to have homogeneous preferences, the fact that they earn different wages causes them to differ in their demand for public goods. The first order condition in (2.25) still holds. However, as a

⁹ Various strategies have been used to calculate a single rental rate for each metro area. For example, Roback (1982) uses the average price of housing per square foot, Blomquist et al. (1988) use average monthly housing expenditures without adjusting for variation in structural characteristics, and Bayer et al. (2006) use data on a vector of structural characteristics for individual homes to estimate the effective “price of housing services” in each metro area. The logic that motivates Bayer et al’s strategy is discussed in part IV of this chapter in the context of sorting models.

result of the complementarity between job skills, the wage rate paid to one type of worker depends on the other type's demand for the public good. This poses an intractable endogeneity problem for the second stage of Rosen's procedure.

(D) *Summarizing the Recent Developments and Remaining Limitations*

Sherwin Rosen's (1974) paper has been extremely influential. It currently has over 1300 citations in the *Social Science Citation Index*, roughly a third of which are in journals dedicated to public, urban, and environmental issues. However, empirical studies are typically limited to the first stage of Rosen's strategy—estimating the hedonic price function to recover measures for the marginal willingness-to-pay. Very few studies implement his second stage to estimate the demand for a public good. The barrier to reduced-form estimation has been the need for data from multiple markets on housing transactions, household demographics, and instruments. The traditional concern with structural estimation is that restrictions on preferences are inherently arbitrary.

The recent contributions by Ekeland, Heckman, and Nesheim (2004) and Bajari and Benkard (2005) have reduced the barriers to second-stage estimation. Ekeland et al. provide some preliminary evidence that instruments and data from multiple markets may be unnecessary to implement the reduced-form approach. Bajari and Benkard start to relax the rigidity of the structural approach by allowing individual households to differ in their tastes for each characteristic, and allowing the choice set to be discrete in some or all of the characteristics. Despite these advances, there are still two major caveats to using the existing hedonic methods for applied welfare analysis: they do not allow general equilibrium adjustment and they do not simultaneously allow preference heterogeneity and labor market choices.

Empirical applications of Roback's (1982) interregional model have indicated that wages reflect a substantial share of the total implicit price of public goods. Yet, at a theoretical level, the interregional model is disconnected from the property value approach. In particular, they rely on contradictory assumptions about mobility and household heterogeneity. Using a metropolitan area as the frame of reference, the property value hedonic identifies preferences from location choices at the intensive margin, while the

interregional model identifies preferences from location choices at the extensive margin. In a setting with homogeneous households, zero moving costs, and a continuous choice set, it would seem natural for there to be a “law of one implicit price function” that would require both models to return the same information about the demand for a public good. It is less clear what to expect in a more realistic setting where households have heterogeneous preferences and skills, some households are retired, moving costs are substantial, and some public goods cannot be continuously chosen.

Bayer, Keohane, and Timmins (2006) relax the national free mobility assumption. They find that indirectly controlling for the cost of moving between metropolitan areas triples their estimates for the marginal willingness-to-pay for an air quality improvement. While their work begins to bridge the gap between the property value and interregional perspectives, unifying the two models still requires reconciling their depictions of household heterogeneity. As Roback (1988) illustrates, introducing heterogeneity into the interregional model makes it difficult to recover the demand for a public good. Section VI returns to this problem in a structural framework.

The second caveat to the hedonic method is that its marginal perspective is only consistent with a partial equilibrium setting. By failing to incorporate the implications of general equilibrium adjustment, hedonic models are incapable of consistently analyzing public policies that would lead to large scale changes in public goods. That is, for a sufficiently large shock to a public good, households may respond by changing their location, which will affect the equilibrium housing prices and feed back into welfare measures. Scotchmer (1986) demonstrated that the information contained in marginal implicit prices is insufficient to predict the pattern of general equilibrium adjustment. Subsequently, Bartik (1988) and Kanemoto (1988) investigated the possibility of using the information in marginal implicit prices to calculate ex ante bounds on general equilibrium welfare measures. While they both report positive results, their conclusions are contradictory. The reason is that they differ in how they define the initial equilibrium and the possibilities for general equilibrium adjustment. Under a very restrictive set of conditions, Kanemoto proves that ex ante welfare measures will overestimate the benefits from an amenity improvement. The requirements for his proof include homogeneous households, only two types of houses, and that the

improvement is funded from tax revenue. These restrictions are relaxed in Bartik's model. He argues that ex ante welfare measures will underestimate the benefits from an (exogenous) amenity improvement when heterogeneous households face a diverse set of housing opportunities. Contrasting Bartik (1988) and Kanemoto (1988) illustrates how assumptions made by the researcher can influence conclusions about the difference between partial and general equilibrium welfare measures.

The two caveats to the hedonic method can be addressed by a recently-developed class of empirical sorting models. Like the property value hedonic, they envision heterogeneous households choosing utility-maximizing locations. The difference is that they divide the housing market into a discrete set of possible location choices. Recent applications have used this framework to develop a theoretically consistent simulation-based approach to general equilibrium analysis. An advantage of the simulation approach is that it allows the analyst to test the sensitivity of results to maintained assumptions about the nature of preference heterogeneity and the set of choice alternatives. Sorting models also offer the potential to develop welfare measures that are consistent with heterogeneous households making simultaneous job and house choices.

IV. Sorting Models

In his seminal 1956 paper, Charles Tiebout suggested a conceptual solution to the problem of providing local public goods efficiently. He envisioned freely mobile households sorting across a finite set of communities based on their preferences for the local public goods provided by those communities. Households that enjoy playing golf will tend to migrate to communities with golf courses and they will tend to vote to spend tax revenue to maintain those golf courses. Tiebout reasoned that, with free mobility, the location choices that households make should reveal their preferences for public goods. In his own words: "*There is no way in which the consumer can avoid revealing his preferences in a spatial economy. Spatial mobility provides the local public-goods counterpart to the private market's shopping trip.*"

The theoretical literature that followed Tiebout's work focused on formalizing his

conceptual model and proving the existence of a “locational equilibrium” in which no household would be better off by moving. Part (A) of this section outlines the theoretical literature by summarizing a series of papers that develop increasingly general depictions of sorting equilibria, supported by existence proofs. For the most part, this line of work treats households as having identical preferences for public goods, and therefore fails to capture a key feature of Tiebout’s reasoning.

Incorporating preference heterogeneity into a theoretical sorting model makes it difficult to develop a general proof supporting the existence of a locational equilibrium. Nevertheless, the empirical literature has moved forward by developing a suite of structural models that start by assuming a locational equilibrium exists and then use the properties of that equilibrium to infer households’ heterogeneous preferences for public goods. As in the structural hedonic literature, recovering preference parameters enables the econometrician to calculate the demand for a public good. Moreover, because sorting models are estimated in a way that characterizes equilibrium in the entire housing market, the estimation results can be used to simulate general equilibrium responses to large scale changes in the public goods.

Parts (B) and (C) of this section describe the two predominant empirical frameworks. Part (B) describes an approach developed by Epple and Sieg (1999). In their model, households maximize utility by choosing a discrete community and a continuous quantity of housing within that community. Their approach is described in detail because some of the elements of their structural model form the basis for the estimator developed in Chapter 3. Part (C) describes a random-utility framework that Bayer, McMillan, and Reuben (2005) recently developed.

(A) Theory

Tiebout’s locational sorting model assumes, *ceteris paribus*, heterogeneous households select a community based on its local public goods. Suppose the housing market can be divided into a finite set of $j = 1, \dots, J$ communities, such that provision of public goods is constant within each community and varies between communities: g_1, \dots, g_J . Figure 2.2C illustrates this depiction of the urban landscape. Each household chooses the community that

maximizes its utility, given its exogenous income and its preferences. For heuristic purposes, utility maximization can be depicted as a two-stage problem where each household first determines the optimal quantities of housing and numeraire in every community and then chooses the community that maximizes its utility. The first stage is shown in (2.27).

$$(2.27) \quad \max_{\bar{h}, b} U(b, \bar{h}, g_j, \alpha) \text{ subject to } y = b + p_j \cdot \bar{h}.$$

Theoretical sorting models treat housing as a homogeneous commodity that can be consumed in continuous quantities at a constant price. This is represented in (2.27) by defining p_j as the annualized per-unit price of housing in community j and \bar{h} as the quantity of housing consumed. Note that the bar superscript on \bar{h} does not imply the quantity of housing consumed is fixed; the quantity consumed will vary across households according to their income, preferences, and chosen community. The bar superscript is meant to differentiate the concept of a homogeneous unit of housing from h , which continues to represent a vector of structural housing characteristics (e.g. bedrooms, bathrooms, sqft.). Part (B) of this section describes the relationship between these two concepts in a particular empirical model.

Assume that zoning does not constrain housing construction. Then households can purchase any quantity of housing at the market price in each community, in which case housing is “optimized out” of the problem and preferences can be restated using the indirect utility function in (2.28).

$$(2.28) \quad V(g, p, \alpha, y) = U[g, \bar{h}(g, p, \alpha, y), y - p \cdot \bar{h}(g, p, \alpha, y), \alpha].$$

Assuming households are price-takers and can move freely between communities, each household will choose the community that maximizes its well-being, given income and prices. An intercommunity *locational equilibrium* is achieved when every household has chosen its utility-maximizing community and nobody wants to move, given housing price and local public goods.

Most of the theoretical work on sorting has focused on characterizing restrictions on preferences that support the existence of a locational equilibrium. Ellickson (1971) initiated

this line of work. Three features of his model form the basis for most of the subsequent literature. First, he assumed that provision of public goods in community j could be represented by a 1-dimensional measure, \bar{g}_j , an index that represents the composite quality of public goods in that community. Second, he assumed that households have homogeneous preferences ($\alpha_i = \dots = \alpha_j$) and therefore differ only in their income. Given the first two assumptions, he imposed the restriction that indifference curves in the (\bar{g}, p) plane are strictly increasing in income. This, he reasoned, would support a locational equilibrium in which households are perfectly stratified across communities by income. Figure 2.5 uses a two-community example to illustrate the idea. Household i is exactly indifferent between the two communities. Any household with lower (higher) income, such as household $i - 1$ ($i + 1$), will always prefer the cheaper (more expensive) community because indifference curves cannot cross more than once.

Using the three restrictions from Ellickson's paper, Westoff (1977) proved that a locational equilibrium exists in a model where households in each community vote to determine public goods provision and community-specific tax rates. Epple, Filimon, and Romer (1984, 1993) extended Westoff's model to include a market for housing that must clear within each community. Finally, Epple and Romer (1991) generalized the model further to allow voters to anticipate the consequences of their votes for housing prices and migration. While these models help to formalize Tiebout's theory, they do a poor job of predicting actual sorting behavior. To see why, let the J communities be ordered by their quality of public goods provision: $\bar{g}_1 < \bar{g}_2 < \dots < \bar{g}_J$. The key restriction in the theoretical literature—Ellickson's single-crossing restriction—implies that households are partitioned across communities by income, as illustrated in figure 2.6. In the figure, every household in community 1 has lower income than every household in community 2, and so on. Of course in reality community-specific income distributions overlap substantially.

One explanation for why households do not perfectly stratify by income is that they differ in their tastes for public goods. Recognizing this, Epple and Platt (1998) extended the Epple-Romer model to allow households to differ in a single heterogeneous parameter that represents their preferences for composite provision of public goods relative to private goods.

This preference heterogeneity complicates the development of a general analytical proof. Consequently, Epple and Platt rely on a Cobb-Douglas specification for utility and use a numerical example to demonstrate that a locational equilibrium may exist.

Other authors have developed existence proofs for models with preference heterogeneity by removing some of the complicating features of the models developed by Epple and his coauthors. Nechyba (1997) for example is able to prove existence by introducing some discreteness into the housing market. Specifically, he imposes exogenous community boundaries and endows households with fixed quantities of housing. Bayer and Timmins (2005) develop an alternative existence proof that allows households to have preferences for multiple public goods and for the demographic characteristics of their neighbors. Their model does not include taxation and, like Epple and Platt (1988), they rely on a specific functional form for utility. They assume the utility function is additively separable in each public good and that households have idiosyncratic tastes for each individual location. The theoretical work by Epple and Platt (1998) and Bayer and Timmins (2005) supports the two different empirical strategies discussed in parts (B) and (C).¹⁰

(B) Estimation: The CES Utility Approach

Epple and Sieg (1999) is the first illustration of how the properties of a locational equilibrium can be used to recover households' preferences for public goods. Their structural model parallels the theoretical literature. This includes treating housing as a homogeneous commodity that can be consumed at a constant (community-specific) price. While this may seem rather abstract for empirical work, Sieg et al. (2002) show that \bar{h} and p_j are related to the hedonic price function. Specifically, if the structural characteristics of housing enter the utility function in a separable sub-function that is homogeneous of degree 1, expenditures on housing can be factored into the product of a price index and a quantity index:

¹⁰ Ferreyra (2005) recently developed a third estimation framework based on Nechyba's (1997) proof and his subsequent empirical model (Nechyba [1999, 2000]). Ferreyra's estimator is distinguished by the fact that it treats public goods provision endogenously. That is, she selects values for the structural parameters that match predicted and observed expenditures on provision of a public good. However, in order to ensure that her model is computationally feasible, she treats households as having homogeneous preferences for the observable public goods. This limits its applicability to the problem being addressed in this chapter.

$P(g, h) = p_j(g) \cdot \bar{h}(h)$. Assuming this restriction holds, the standard (reduced-form) hedonic price function can be used to recover a consistent index for the price of housing¹¹.

Meanwhile, \bar{h} is optimized out of the problem under the assumption that households are free to choose continuous levels of the structural characteristics of housing, no matter where they choose to live. In other words, Epple and Sieg (1999) depict a discrete-continuous optimization problem where households choose one of a discrete number of communities and then, conditional on that selection, a continuous quantity of housing within that community.

Their econometric strategy is to assume that an equilibrium exists and then use the features of that equilibrium to guide the estimation. As in the theoretical literature, the locational equilibrium is characterized by a single-crossing restriction. To formalize the restriction, equation (2.29) shows the slope of an “indirect indifference curve” in (\bar{g}, p) space.

$$(2.29) \quad M(\bar{g}, p, \alpha, y) = \left\langle \frac{dp}{d\bar{g}} \Big|_{V = \bar{V}} \right\rangle = - \frac{\partial V(\bar{g}, p, \alpha, y) / \partial \bar{g}}{\partial V(\bar{g}, p, \alpha, y) / \partial p}.$$

Assuming M is monotonically increasing in $(y | \alpha)$ and $(\alpha | y)$, indifference curves in the (\bar{g}, p) plane satisfy single crossing. This has an intuitive interpretation. Roy’s Identity implies that $-\partial V(\cdot) / \partial p$ must equal the marginal utility of income, $\lambda = \partial V(\cdot) / \partial y$, times the Marshallian demand for housing, $\bar{h}(\bar{g}, p, \alpha, y)$.

$$(2.30) \quad M(\cdot) = - \frac{\partial V(\cdot) / \partial \bar{g}}{\partial V(\cdot) / \partial p} = \frac{\partial V(\cdot) / \partial \bar{g}}{\lambda h(\cdot)} = \frac{1}{h(\cdot)} \left[\frac{\partial V(\cdot) / \partial \bar{g}}{\partial V(\cdot) / \partial y} \right].$$

The term in brackets in equation (2.30) is the Marshallian virtual price of public goods. Therefore, the single crossing restriction implies that the Marshallian virtual price, per unit of housing, is strictly increasing in income and in preferences for public goods relative to

¹¹ This process is described in detail as part of the description of the estimator in chapter 3.

private goods¹².

The single crossing property implies that, in equilibrium, three properties characterize sorting by each household type: *boundary indifference*, *stratification*, and *increasing bundles*. Without loss of generality, let the J locations be ordered according to the index of public goods, $\bar{g}_1 < \dots < \bar{g}_J$. *Boundary indifference* requires a household on the “border” between two locations in (α, y) space to be exactly indifferent between those locations. Equation (2.31) defines the set of border individuals. It must hold for all $j = 1, \dots, J - 1$.

$$(2.31) \quad \left\{ (\alpha, y) : V(\bar{g}_j, p_j, \alpha, y) = V(\bar{g}_{j+1}, p_{j+1}, \alpha, y) \right\}.$$

The *increasing bundles* property requires that for any two locations in the ordering, $(j, j + 1)$ equation (2.32) must hold.

$$(2.32) \quad y_{j+1}(\alpha) > y_j(\alpha) \Rightarrow p_{j+1} > p_j \quad \text{and} \quad \bar{g}_{j+1} > \bar{g}_j.$$

That is, the ranking of communities by public goods provision must match the ranking by price. The third property, *stratification*, requires that households of each type are stratified across the J ordered locations by $(\alpha | y)$ and by $(y | \alpha)$, as defined in (2.33).

$$(2.33) \quad \begin{aligned} & (y_{j-1} | \alpha) < (y_j | \alpha) < (y_{j+1} | \alpha) \\ & \text{and} \\ & (\alpha_{j-1} | y) < (\alpha_j | y) < (\alpha_{j+1} | y) \end{aligned}.$$

Figure 2.7 illustrates the implied partition of households into communities in (y, α) space. Conditional on preferences, higher income households always choose to live in communities with more public goods. Likewise, conditional on income, households with stronger preferences always choose communities with more public goods.

¹² This property is related to the Willig condition that is often applied together with weak complementarity to identify the Hicksian willingness to pay for changes in public goods. The Willig condition requires the willingness-to-pay per unit of the weak complement to be constant at all levels of income. See Smith and Banzhaf (2004) or Palmquist (2005) for details.

While the three properties are necessary for a locational equilibrium to exist, they are not sufficient. Any locational equilibrium must also be characterized by a vector of housing prices and a vector of public goods such that no household could increase its utility by moving. Epple and Sieg's strategy is to assume that an equilibrium exists and focus on recovering values for the structural parameters that justify observed (equilibrium) location choices.

Equation (2.34) shows their CES indirect utility function for household i in community j . The first term represents utility from public goods. \bar{g} is a linear index of public goods, all but one of which are observable. Notice that the weights in the index are constant. This requires households to agree on the ranking of communities by \bar{g} . However, households can differ in the strength of their preferences for public goods relative to private goods through the α_i term.

$$(2.34) \quad V_i = \left\{ \alpha_i (\bar{g}_j)^\rho + \left[\exp\left(\frac{y_i^{1-\nu} - 1}{1-\nu}\right) \exp\left(-\frac{\beta(p_j^{\eta+1}) - 1}{1+\eta}\right) \right]^\rho \right\}^{\frac{1}{\rho}},$$

$$\text{where } \bar{g}_j = \gamma_1 g_{1,j} + \gamma_2 g_{2,j} \dots + \gamma_{N-1} g_{N-1,j} + \gamma_N \bar{\xi}_j, \quad F(\alpha, y) \sim \text{lognormal}.$$

The second term in the expression represents utility from private goods. The price and income elasticities of the demand for housing are represented by η and ν , β is a housing demand intercept, and ρ is the elasticity of substitution between public and private goods. The joint distribution of α_i and income is assumed to be lognormal.

In their application to the Boston Metropolitan Area, Epple and Sieg define each community as a school district. Then, using data on housing prices, household income, and observable public goods in 92 school districts, they estimate the parameters of the model in two stages. First, they estimate a subset of the parameters by minimizing the distance between predicted and observed quantiles of community-specific income distributions. Then, the second stage of their estimator uses the boundary indifference and increasing bundles properties, along with the lognormality assumption, to solve for the remaining

structural parameters, including an unobserved public good in each community $(\bar{\xi}_1, \dots, \bar{\xi}_J)$. As in the discrete-choice hedonic model, $\bar{\xi}_j = f(\xi_j)$ can be interpreted as an index of all the unobserved public goods in community j .

Sieg, Smith, Banzhaf, and Walsh (2004) refined the estimator by adding moment conditions based on the distribution of housing prices and using a GMM approach to estimate all the parameters of the model simultaneously. They also demonstrated that the model can be used to calculate general equilibrium welfare effects from large-scale changes in public goods. Their approach uses the parameter estimates to simulate how households would respond to a change. Intuitively, a large-scale change in the spatial distribution of public goods may induce some households to move, changing equilibrium housing prices which, in turn, would feed back into welfare measures. They use the model to estimate preferences for air quality in 93 school districts in the Los Angeles Air Basin. In a companion paper, Smith, Sieg, Banzhaf, and Walsh (2004) calculate general equilibrium welfare measures from the ozone improvements predicted for the Los Angeles air basin as a result of the Clean Air Act Amendments.

One of the key limitations in this set of papers is the assumption that all households agree on the ranking of communities by public goods provision. Relaxing this restriction is important because it is reasonable to expect households to evaluate component of a vector of local public goods quite differently. For example, households with school age children may be more concerned about school quality while retirees may place more emphasis on climate and other environmental amenities. The random utility specification in part (C) allows this flexibility.

(C) *Estimation: The Random Utility Approach*

Bayer, McMillan, and Reuben. (2005) develop an alternative sorting model that combines elements of the hedonic and sorting literatures. Like the hedonic property value models, they define the object of choice as an individual house which is defined by a vector of structural characteristics. Then, like the sorting literature, they assume that each house is located within a community and that public goods are homogeneous within communities. One of the

distinguishing features of their model is its extremely flexible treatment of preference heterogeneity. Households differ in their preferences for: (1) each public good; (2) each structural characteristic of housing; and (3) each demographic characteristic describing other households in their community. In addition, each household has an idiosyncratic “taste” for every individual house. With this specification, the number of random parameters in the model exceeds the number of households times the number of houses.

Equation (2.35) illustrates their indirect utility function. In a change of notation (limited to this equation) the superscripts index households and the subscripts on the heterogeneous preference parameters $[\alpha_h^i, \alpha_g^i, \alpha_d^i, \alpha_p^i]$ index the partition of preference parameters into a set of vectors that correspond to preferences for structural characteristics of housing, public goods, demographics, and the price of housing.

$$(2.35) \quad V_{i,j} = \alpha_h^i h_j + \alpha_g^i g_j + \alpha_d^i d_j + \alpha_p^i P_j + \bar{\xi}_j + \varepsilon_j^i,$$

$$\text{where } \alpha_a^i = \alpha_{0,a} + \sum_{r=1}^R \alpha_{r,a} d_r^i, \quad \text{and } \varepsilon_j^i \sim \text{iid type I extreme value.}$$

Each of the structural preference parameters can be decomposed into a common component, α_0 , and a set of terms that interact with the household’s demographic characteristics, d^i . Income is included in the vector of demographic characteristics. Finally, the idiosyncratic tastes are assumed to follow the standard logit distributional assumptions.

Because idiosyncratic tastes follow the type I extreme value distribution, there is a positive probability that every household chooses every house. Households do not “locate” in houses. There is a set of probabilities they will locate in each, and the probability is one that each house is occupied. This condition (supply = expected demand) allows them to express the probability of choosing a house as a function of the structural parameters. They use a version of the estimator from Berry, Levinsohn and Pakes (1995) to find values for the heterogeneous preference parameters that minimize the difference between predicted and observed location choices. The residuals in the objective function are treated as estimates of the house-specific unobserved characteristics $(\bar{\xi}_1, \dots, \bar{\xi}_j)$. The flexibility of the random utility specification allows Bayer et al. to recover detailed information on the marginal willingness-

to-pay for changes in public goods.

In addition to calculating MWTP, Bayer et al. demonstrate that their model can be used to simulate general equilibrium adjustments to a large-scale change in public goods. However, it is unclear whether welfare measures derived from their model will be theoretically consistent. There are two concerns. First, because location choices are probabilistic, so are general equilibrium adjustments. This makes it difficult to interpret the resulting welfare measures in the context of existing theory (Scotchmer [1986]). The second concern is that it seems unlikely the indirect utility function in (2.35) can satisfy homogeneity of degree zero in prices and income. Usually the homogeneity requirement is satisfied by implicitly normalizing housing prices and income by the price of the numeraire good. The budget constant in (2.27) provides an example. Following this logic, the budget constraint and direct utility function implied by Bayer et al.'s model can be combined to produce the following theoretically consistent expression for indirect utility:

$V = U(g, h, d, y - P_j, \alpha)$. The problem with the parameterization of this expression in (2.35) is that it depicts y as interacting with each of the housing characteristics while P_j is treated as a housing characteristic. It is difficult to imagine how this can be reconciled with the theoretically consistent general expression.

V. Unifying the Structural Hedonic and Sorting Models

Hedonic and sorting models are closely related. They both depict a frictionless market characterized by perfect information, free mobility, and consumers who act according to their heterogeneous income and preferences. As a result, the two models make similar predictions. Tiebout and Rosen both recognized this. In the special case with no economies of scale in producing public goods, Tiebout (1956) suggests that households will choose communities that *exactly* match their preferences, effectively making each household its own

local government¹³. This reasoning anticipates the hedonic first-order conditions. Likewise, hedonic theory predicts that consumers with similar income and preferences will purchase products with similar characteristics. Rosen (1974) notes the similarity between this (private) market segmentation and the segmented market for public goods in Tiebout's work¹⁴. Since both models start from a similar premise and, through sorting behavior, achieve similar results it seems natural to ask: Are hedonic and sorting models isomorphic?

Their theoretical models differ in how they portray supply. In the hedonic model atomistic sellers offer a continuum of choices, while sorting theory envisions a discrete set of communities which tax themselves to produce public goods. However, the empirical applications of the two models discussed in sections III and IV abstract from the production process and treat the available choices as exogenous. This leaves discreteness/continuity of the choice set as the key difference between them. As a result, it should come as no surprise that Bajari and Benkard's (2005) discrete-choice hedonic model *is* isomorphic to the sorting models developed by Epple and Sieg (1999) and Bayer, McMillan and Reuben (2005)—at least at a conceptual level. All the models take prices as given, assume that observed choices reflect utility-maximizing behavior, and then concentrate on recovering values for the structural parameters that explain those choices. The differences in their econometric methods stem from how they choose to represent the structure of the utility function and preference heterogeneity—decisions which are not necessarily tied to the underlying theory.

To compare the depiction of preference heterogeneity in the three models (henceforth ES, BB, and BMR) consider an example where communities differ in two observable public goods (g_1, g_2). In all three models the indirect utility function contains a separable, linear sub-function of public goods which is equivalent to a linear index of overall public goods

¹³ From page 421 of Tiebout (1956), "...the consumer-voters will move to that community which *exactly* satisfies their preferences. ...this may reduce the solution of the problem to the trite one of making each person his own municipal government."

¹⁴ From page 40 of Rosen (1974), "...a clear consequence of the model is that there are natural tendencies toward market segmentation, in the sense that consumers with similar value functions purchase products with similar specifications. ...In fact, the above specification is very similar in spirit to Tiebout's (1956) analysis of the implicit market for neighborhoods, local public goods being the 'characteristics' in this case."

provision. Let it be represented by $\bar{g}(\gamma)$, where γ is a vector of relative preferences or index “weights”. Using this notation, the three indices are represented as follows:

$$(2.36) \quad \begin{array}{ll} \text{(ES)} & \bar{g}_j = \gamma_1 g_{1,j} + \gamma_2 g_{2,j} + \bar{\xi}_j \\ \text{(BB)} & \bar{g}_j = \gamma_{i,1} g_{1,j} + \gamma_{i,2} g_{2,j} + \bar{\xi}_j \quad \text{where} \quad \bar{\xi}_j = f(\xi_j, \omega) \\ \text{(BMR)} & \bar{g}_j = \gamma_{i,1} g_{1,j} + \gamma_{i,2} g_{2,j} + \bar{\xi}_j + \varepsilon_{i,j} \quad \varepsilon_{i,j} = f(\xi_j, \varphi_i). \end{array}$$

$\bar{\xi}_j$ and $\varepsilon_{i,j}$ are both indices of all the unobserved public goods at location j . The difference between them is that $\bar{\xi}_j$ is characterized by a vector of constant weights (ω) while $\varepsilon_{i,j}$ is characterized by a vector of individual-specific weights (φ_i).

In the ES specification, all the weights are constant. In other words every household is required to have the same relative preferences for every public good, observed and unobserved. Everybody trades air quality and school quality at the same rate for example. The BB specification relaxes this depiction of preference heterogeneity by allowing households to differ in their relative preferences for the observed public goods. However, households still have the same relative preferences for each pair of unobserved public goods. The BMR specification relaxes this restriction by introducing idiosyncratic tastes for each location. This effectively allows each household to differ in its relative preferences for each unobserved public good, even though the $\varepsilon_{i,j}$ term is only one-dimensional.

The depiction of preference heterogeneity determines the way in which households perceive communities to be differentiated. Suppose that households differ only in their income and preferences for public goods¹⁵. When households share the same relative preferences for each pair of public goods in the ES specification, every household must also agree on a common ranking of communities by the quality of the public goods they provide. This ranking is based on the \bar{g} index in (2.36). The notion that a set of differentiated communities can be unanimously ranked by quality is analogous to Lancaster’s (1979)

¹⁵ This is consistent with the maintained assumptions of the ES model that housing can be treated as a homogeneous commodity that can be consumed in continuous quantities, and that all households share the same values for the price and income elasticities of housing.

description of “vertical” product differentiation. In contrast, when households differ in their relative preferences for public goods, as in BB and BMR, they may also differ in the way they rank communities—a situation analogous to Lancaster’s description of “horizontal” product differentiation. The distinction between vertical and horizontal differentiation is important because the two concepts differ in their implications for substitution patterns (Anderson, de Palma, and Thisse [1992]). When households agree on the ranking of communities by overall public goods provision in the vertically differentiated case, they must also agree on the opportunities for spatial substitution. For example, if asked to identify the two closest substitutes for community j , every household in the ES specification will select the two adjacent communities in the ranking by \bar{g} . Horizontal differentiation allows more diversity in perceived substitution possibilities. Since households in the BB and BMR specifications can differ in how they rank communities by \bar{g} , they may disagree on which communities represent the closest substitutes for j .

The rest of this section explores how the different representations of preference heterogeneity used by ES, BB, and BMR influence the substitution patterns and welfare measures predicted by their models. Part (A) demonstrates that all three models use the same fundamental revealed preference logic to identify heterogeneous preference parameters, and therefore welfare measures. The identification stems from two sources: structural restrictions on the indirect utility function and maintained assumptions about the distribution of preferences. Part (B) illustrates how those same structural restrictions and distributional assumptions effectively determine the predicted substitution patterns between locations. Part (C) starts by explaining why the analyst must account for unobserved public goods. ES, BB, and BMR each use a different strategy to identify unobserved public goods, and again the identification is motivated by distributional assumptions. Finally, part (D) concludes by returning to the question of whether hedonic and sorting models are isomorphic. Throughout this discussion, the “vertical” and “horizontal” terminology is used to distinguish between representations of preference heterogeneity that imply households have the same ranking of communities (i.e. vertical) and representations that imply a diversity of rankings (i.e. horizontal).

(A) *Identifying Heterogeneous Preferences from Structural Restrictions*

When a consumer purchases a bundle of goods, the composition of that bundle reveals some basic features of his preferences. He could have chosen any other bundle within his budget. The fact that he did not reveals that he considers them to be inferior. Samuelson (1948) demonstrated that comparing bundles that were purchased with bundles that could have been purchased, but were not, can identify bounds on the consumer's indifference curves.

To consider how Samuelson's revealed preference logic underlies the structural hedonic and sorting models, suppose households maximize their utility by sorting among the four communities described in table 2.4. Assume a household chooses to live in community 3. Given preferences that are monotonic and quasiconvex, the choice not to live in community 2 identifies bounds on the shape of that household's indifference curves. Figure 2.8 illustrates this graphically, using the technique from Varian (1982). The two points show the locations of communities 2 and 3 in characteristics space. The shaded areas bound the set of indifference curves passing through the point C_2 that are consistent with a household that chooses to live in community 3. This result requires nothing more than monotonicity, quasiconvexity, and the assumption of utility maximization¹⁶. Repeating this exercise to compare community 3 with communities 1 and 4 provides two more sets of bounds on the household's indifference curves. To point identify a household's indifference curve within these bounds requires imposing some structure on the utility function.

Two types of structural restrictions are required to point-identify households' preferences based on their observed location choices. First, a parametric indirect utility function must be selected. Second, a distribution must be specified for each preference parameter in that function used to characterize household heterogeneity. Each restriction makes a different type of contribution to the identification.

Distributional assumptions are necessary due to the discreteness in the choice set. When household i chooses j from a finite set of communities, utility maximization is characterized by the set of inequalities in equation (2.37).

¹⁶ In chapter 8 of his *Microeconomic Analysis* textbook, Varian provides an intuitive derivation of the shaded areas. In short, the horizontal and vertical lines are implied by monotonicity and the diagonal line by quasiconvexity.

$$(2.37) \quad V[\bar{g}_{i,j}(\gamma_i), p_j, \alpha_i, y_i] \geq V[\bar{g}_{i,k}(\gamma_i), p_k, \alpha_i, y_i], \quad \forall k = 1, \dots, J.$$

Given a parametric form for the indirect utility function, the inequalities provide set identification of the heterogeneous preference parameters. It must be the case that $(\alpha_i, \gamma_i) \in A_{i,j}$, where $A_{i,j} = \{ (\alpha_i, \gamma_i) : (\alpha_i, \gamma_i) \text{ satisfies (2.37)} \}$. In words, the choice of community j reveals only that household i 's preferences lie somewhere in the $A_{i,j}$ set. Imposing a distribution on (α, γ) allows the analyst to identify the density of preferences within $A_{i,j}$.

To illustrate the role of each type of restriction in identifying preferences, consider a specific example using the following CES indirect utility function¹⁷:

$$(2.38) \quad V[\bar{g}(\gamma), p, \alpha, y] = \left\{ \alpha_i (\bar{g}_{i,j})^{-0.01} + .79 [\exp(4y_i^{.25} - 54p_j^{.04})]^{-0.01} \right\}^{-0.01},$$

where $\bar{g}_{i,j} = \gamma_{i,air} AIR_j + \gamma_{i,school} SCHOOL_j$.

The first term represents utility from public goods, and the second term represents utility from the private good component of housing. Households differ in their income and in their preferences for a linear index of two public goods that differentiate communities, air quality and school quality. (Unobserved public goods are suppressed for now to keep the example relatively simple.) There are two components of preference heterogeneity. Households differ in the relative weights they place on each public good in the index $(\gamma_{i,air}, \gamma_{i,school})$ and in the overall strength of their preferences for public goods relative to private goods (α_i) . This depiction of preference heterogeneity is useful for illustrating the differences between vertical and horizontal differentiation as depicted in ES and BB.

To illustrate how this indirect utility function provides set identification of preferences, first consider the simplest form of preference heterogeneity—*vertical*

¹⁷ This CES function provides the basis for the subsequent structural model. Specifically, it is the indirect utility function from the empirical model in Chapter 3 with $\beta = 2$, $\eta = -.963$, $\nu = .75$, and $\rho = -.01$. If the weights in the public goods index are constant, it reduces to the indirect utility function in Epple and Sieg (1999).

differentiation. In a vertically differentiated model, all the variation in tastes can be condensed into a single heterogeneous preference parameter that ranks locations by “quality”. The CES utility function simplifies to this case when households are constrained to have the same relative preferences for the two public goods. For example, let the weights be: $(\gamma_{i,air}, \gamma_{i,school}) \equiv (0.48, 0.52), \forall i$. With constant weights, all households agree on a common ranking of communities by the public goods index, and sort according to their income and α_i . By conditioning on income, the system in (2.37) can be solved for the bounds of the α_i sets that rationalize each location choice using the specific parameter values given in (2.38). Figure 2.9 displays the partition of α at $y=\$50,000$.

The figure illustrates two critical limitations of set identification. First, preferences are not point identified within the bounds of a set. The choice of community 2 reveals only that the household’s preferences lie somewhere in $A_{i,2}$: $1.09 \leq \alpha_i \leq 1.28$. A second important limitation of set identification is that the preference set that corresponds to the highest (lowest) provision of public goods is not bounded from above (below) by the revealed preference logic in (2.37). These two limitations require that a distribution be specified for α_i . This added information transforms the observed location choices by a population of households into a distribution of preferences.

When vertical differentiation is relaxed, observed location choices are required to set-identify more heterogeneous preference parameters. Returning to the CES example, horizontally differentiated preferences imply households differ in their relative preferences for the two public goods. This generalization complicates the partitioning process. For example, at $y=\$50,000$ the inequalities in (2.37) can be solved for a large number of $(\alpha, \gamma_{air}, \gamma_{school})$ points that would make a household exactly indifferent between each pair of communities. Figure 2.10A shows the “indifference loci” formed from these points in \mathfrak{R}_+^2 . A household with preferences anywhere along the locus from (0.1, 0) to (4, 0.49) would be exactly indifferent between communities 2 and 3. Households above (below) this line have stronger (weaker) relative preferences for air quality, and therefore prefer community 2 (3). The indifference loci delineate 12 areas, each of which corresponds to a different ranking of

communities by utility. The rankings (from highest to lowest) are shown in the figure. Notice that adjacent areas often have the same utility maximizing community. Community 1 maximizes utility throughout the region defined by $[(0,0), (0,1), (0.87,1), (1,0.47), (0.87,0)]$, for example. By taking the union of areas that share the same utility-maximizing community, figure 2.10B partitions preference space into regions that rationalize each of the four community choices. For example, taking the union of the two areas that have community 1 as the top-ranked choice (the areas in figure 2.10A with rankings: 1,2,3,4 and 1,3,2,4) produces region $A_{i,1}$ in figure 2.10B. A household with preference anywhere in region $A_{i,1}$ will maximize its utility by choosing to live in community 1.

This partitioning process illustrates how the identifying power of the indirect utility function differs under vertical and horizontal differentiation. In the vertical case the choice of community 2 indicates that the household's preferences belong to the set: $(\gamma_{air} = .48, 1.09 \leq \alpha \leq 1.28)$, which appears in figure 2.10B as the dashed line in the lower left corner of the $A_{i,2}$ region. $A_{i,2}$ is the preference set identified by the choice of community 2 in the horizontal case. This comparison illustrates a general principle: *preference sets revealed by vertically differentiated sorting are subsets of their horizontally differentiated counterparts.*

In an empirical analysis, distributional assumptions will influence estimated welfare measures for policy changes. The marginal-willingness-to-pay (MWTP) for public goods is a function of $(\alpha, \gamma_{air}, \gamma_{school})$ which means every point in figure 2.10B corresponds to a specific MWTP. Suppose we want to infer the distribution of MWTP for air quality for households living in community 3. The choice of community 3 reveals only that households living there have preferences somewhere in $A_{i,3}$. To calculate their distribution of MWTP we must first specify a function for the preference parameters over the $A_{i,3}$ region. Two extreme cases provide bounds for the MWTP. The first case is where every household has preferences at the point (*), which corresponds to the lowest MWTP of any point in $A_{i,3}$. The opposite extreme is where every household has preferences at (**), which corresponds to the highest MWTP. Thus, $[MWTP(*), MWTP(**)]$ spans the range of possible measures for

individual MWTP. The wider this range the greater the sensitivity of welfare effects to the distributional assumptions made in order to move from set to point identification.

The figure also illustrates there are limits to what can be learned from revealed preference analysis. Consider community 4. Because it provides the most public goods, it will attract households with the strongest preferences and the highest MWTP. These households may make the largest contribution to summary measures of the average MWTP. In this case, revealed preference analysis is limited because there is no upper bound on α_i in region $A_{i,4}$ of the partition. To recover a MWTP distribution for community 4, either an absolute upper bound must be imposed on α_i or a distribution that limits weight in the tail.

If the public goods index in (2.38) were generalized to allow idiosyncratic tastes for individual location choices, the partition would expand into four more dimensions: $(\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4)$. While this 6-dimensional partition is difficult to visualize, it will have two obvious properties. First every region will be unbounded. For example, region 2 is clearly unbounded in the ε_2 dimension because increasing $\varepsilon_{i,2}$ always increases the utility from locating in community 2. Given values for the other preference parameters $(\alpha, \gamma_{air}, \varepsilon_1, \varepsilon_3, \varepsilon_4)$ there must be some threshold for ε_2 above which community 2 will be the utility-maximizing location. The second distinguishing feature of the partition is that every region must share a border with every other region. Within region 1 for example, ε_4 can always be increased until communities 1 and 4 provide exactly the same utility, and the same is true for ε_2 and ε_3 .

The vertical/horizontal modeling choice is similar to a bias/variance tradeoff. For example, suppose that horizontal differentiation is the “true” form of preference heterogeneity. By restricting relative preferences, vertical differentiation biases welfare measures. Horizontal differentiation eliminates the restriction that causes bias, but the added dimensionality of preferences increases the scope for distributional assumptions to influence results. Adding idiosyncratic tastes further reduces the potential for bias due to incorrect restrictions but simultaneously increases the variance.

(B) *Identifying Substitution Patterns*

Structural restrictions on the utility function control the scope of substitution patterns. With vertical differentiation each community has at most two substitutes, the adjacent communities in the ranking by public goods¹⁸. With horizontal differentiation the total number of substitutes for each community falls between 2 and J , depending on the number of choices relative to the number of public goods (Anderson, DePalma and Thisse [1992]). The communities that are substitutes will share “borders” in the partition of preference space. Community 2, for example, shares borders with each of the other three communities in figure 2.10B. Consider a marginal increase in the price of housing in community 2. Households that currently reside in 2 but have preferences on the border between 2&4 will respond to the price increase by moving to community 4. Likewise, households on the borders between 2&1 and 2&3 will move to communities 1 and 3.

The substitution patterns that arise from horizontal differentiation have intuitive properties. Locations that are similar in terms of prices and public goods are more likely to be substitutes than those that are not. Notice that in figure 2.10B the two communities with intermediate levels of public goods, 2 and 3, share borders with each of the other three locations while the most and least expensive locations, 1 and 4, do not share a border. Because locations 1 and 4 are furthest removed in terms of prices and public goods, it seems natural to expect that there are few, if any, households that consider them as close substitutes.

When idiosyncratic tastes are added to the horizontally differentiated model every pair of communities will share a border in preference space. This implies that households view every pair of communities as substitutes so that each community has $J-1$ total substitutes. Thus, to summarize, the representation of preference heterogeneity used by BB suggests a set of substitution patterns that falls somewhere between the two different extremes depicted in ES and BMR.

ES, BB, and BMR also differ in how they restrict substitution between public and private goods. All three models assume that different public goods are perfectly

¹⁸ The definition of substitution used here is defined as “strong gross substitution” in Anderson, DePalma and Thisse (1992), where k is a substitute for j iff $\partial h_k / \partial P_j > 0$.

substitutable. This conclusion follows from the additive form of the public goods index. BMR and BB also treat public goods as being perfect substitutes for private goods by virtue of their linear indirect utility functions. In contrast, the CES indirect utility function used by ES allows the relationship between public and private goods to range from perfect substitutes to perfect complements.

(C) *Identifying Unobserved Public Goods and Rationalizing Choices*

The discussion in parts (A) and (B) focused on observed characteristics. In applied work the observed characteristics rarely explain all the observed choices. To illustrate the nature of the problem, a fifth location is added to the choice set in table 2.5. Community 5 is more expensive than community 4 and it has lower levels of both air quality and school quality. Given the specification for utility in equation (2.38) the choice of community 5 cannot be explained without assigning negative preferences for at least one of the observed public goods. Does this necessarily imply that households in community 5 obtain negative marginal utility from public goods? Another possibility is that the high price in community 5 reflects an important unobserved characteristic.

Lancaster (1966) recognized that unobserved characteristics pose a problem for the characteristics approach to consumer theory. He suggested that the analyst continue adding characteristics to the model until every observed choice can be explained. While his reasoning is sound it is also conceivable that in any given application there will be some important characteristics that simply cannot be observed. For example, in their application to computers, Bajari and Benkard (2005) find that even after controlling for an exhaustive set of 19 characteristics, more than half of the computers are still “dominated” like community 5. While the authors of previous sorting applications have not reported the number of dominated locations, the rank violations reported in Sieg et al. (2002) indicate a lower bound of 31%. Thus, the challenge is to develop an econometric strategy that identifies an unobserved location-specific public good which can rationalize observed behavior. Put simply, the estimate for $\bar{\xi}_j$ must “pick up the slack” in explaining choices.

ES, BB, and BMR use different strategies to identify unobserved public goods. In

each case the identification of $\bar{\xi}_j$ is intertwined with the rest of the estimation. However, the general intuition for their identification strategies can be seen by comparing the expressions they use to calculate $\bar{\xi}_j$:

$$\begin{aligned}
 (2.39) \quad & \text{(ES)} \quad \bar{\xi}_j = \gamma_1 g_{1,j} + \gamma_2 g_{2,j} - \delta(\alpha), \quad \text{where } F(\alpha, y) \sim \text{lognormal} \\
 & \text{(BB)} \quad \bar{\xi}_j = F_{p|g=g_j}(p_j), \quad \text{where } \bar{\xi}_j \sim U(0,1) \\
 & \text{(BMR)} \quad \bar{\xi}_j = \gamma_1 g_{1,j} + \gamma_2 g_{2,j} - \delta(g, \gamma_i, \varepsilon_{i,j}), \quad \text{where } \varepsilon_{i,j} \sim \text{type I extreme value.}
 \end{aligned}$$

ES and BMR both treat $\bar{\xi}_j$ as a structural error term. More precisely, $\bar{\xi}_j$ is the residual to a moment condition that is minimized by choosing values for the weights in the public goods index. This approach is intuitive. It maximizes the extent to which observed public goods explain location choices by minimizing the extent to which $\bar{\xi}_j$ is required to “pick up the slack”. While it may not be clear from the equations, this approach guarantees that $\bar{\xi}_j$ will explain observed behavior¹⁹.

Bajari and Benkard develop a completely different approach to recovering $\bar{\xi}_j$ based on hedonic theory. Using a nonparametric approach suggested by Matzkin (2003), they calculate $\bar{\xi}_j$ using information contained in the residual to the reduced form hedonic price function. Intuitively, after controlling for the effect of all the observed characteristics (i.e. structural characteristics of housing and observed public goods) the remaining variation in prices should contain information about the quality of unobserved characteristics associated with each choice. Chapter 3 describes this econometric strategy in greater detail. Like the strategies used by ES and BMR, it guarantees that the estimated values of the unobserved public goods will explain observed choices.

By more fully utilizing the information contained in the hedonic price function, BB’s approach eliminates the reliance on distributional assumptions to identify $\bar{\xi}_j$. This may seem contradictory since each of the expressions for $\bar{\xi}_j$ listed above is associated with a

¹⁹ Recall however that BMR give the concept of location choice a probabilistic interpretation.

distributional assumption. However, the assumption made by BB is innocuous. Since $\bar{\xi}_j$ is an index it does not have any natural units and can be freely normalized by a uniform distribution to a 0-1 scale. In contrast, the assumptions maintained by ES and BMR are arbitrary and are likely to influence the resulting estimates for $\bar{\xi}_j$.

The background and intuition provided here serve to begin the process of highlighting the issues associated with recovering unobserved public goods as a means to explain observed location choices. For example, the consequences of the different identification strategies for the implied welfare measures have yet to be systematically explored. This task is left for future research.

(D) *Moving Toward A Unified Structural Model of Location Choice*

Table 2.6 summarizes the different modeling choices made by ES, BMR, and BB. While they impose different restrictions on preference heterogeneity and substitution patterns, the restrictions are theoretically interchangeable. A vertically differentiated hedonic model would be perfectly consistent with hedonic theory. Likewise, sorting theory does not dictate any particular set of substitution patterns, and so on. The bottom line is that the restrictions in the table can be mixed and matched in any combination to produce an empirical model that could accurately be labeled a “hedonic model” or a “sorting model”.

To take the argument one step further, consider that Bajari and Benkard (2004) prove under fairly general conditions that as the choice set “fills up” the preference sets identified by the discrete choice hedonic will converge to the singletons identified by the continuous choice hedonic. By transitivity, their result reinforces the connection between sorting models and the continuous-choice hedonic:

$$(\textit{sorting}) \approx (\textit{discrete choice hedonic}) \xrightarrow{\textit{as } J \rightarrow \infty} (\textit{continuous choice hedonic}).$$

This flow diagram suggests it might be possible to unify the hedonic and sorting models more formally. Since the two models depict the same consumer behavior, the difficulty in unifying them is likely to stem from their different interpretations of supply. It might be

possible to avoid this issue by applying Bajari and Benkard's (2005) proof that a well-behaved hedonic price function exists without any supply side assumptions. However, the supply side cannot be ignored if the analyst intends to simulate general equilibrium responses to shocks. A formal attempt to unify the hedonic and sorting models is left for future research.

Since the restrictions in table 2.6 are interchangeable there may be gains from recombining them in a way that draws on the individual strengths of the three estimators. To begin, consider the three different sets of maintained assumptions on the properties of the choice set. ES's discrete-continuous treatment is the only one which is necessarily consistent with Samuelson's (1948) underlying revealed preference logic. This is because allowing housing to be purchased in continuous quantities guarantees that a household can locate anywhere without violating its (unobserved) budget constraint. For example, households which cannot afford a 4-bedroom home in Beverly Hills can probably still afford a studio apartment. In contrast, when BMR define the choice set as a discrete collection of individual homes they probably overextend the identifying power of revealed preferences by requiring every household to be financially capable of purchasing every house²⁰.

When it comes to selecting a form of preference heterogeneity, it seems intuitively plausible that households would have horizontal differentiated preferences for the observed public goods. Although as part (A) discussed, allowing preferences to be horizontally differentiated increases the extent to which welfare measures will be sensitive to distributional assumptions. It is also intuitively plausible that households will differ in their preferences for individual locations in a way that the econometrician cannot observe. A household may be attached to a particular location because their family and friends live there for example. This can be addressed by adding idiosyncratic location-specific tastes to the model. However, the extra flexibility comes at a considerable cost because the current class of estimators that can handle that heterogeneity predict location probabilities rather than explain location choices. Given these tradeoffs, the intermediate depiction of preference

²⁰ Banzhaf and Smith (2006) demonstrate that misspecifying the choice set in a random utility model of households' location choices to include homes that lie outside a household's true budget set can have a systematic effect on estimates of the willingness-to-pay for changes in a spatially differentiated amenity.

heterogeneity used by BB seems like the most prudent choice. It will also produce an intuitively plausible set of substitution patterns, as part (B) discussed.

Finally, the approach that BB use to identify the unobserved characteristics has the obvious attraction of more fully exploiting hedonic theory and simultaneously relaxing the need for distributional assumptions. The estimator developed in Chapter 3 combines Bajari and Benkard's strategy for recovering $\bar{\xi}_j$ with their depiction of preference heterogeneity and ES's discrete-continuous depiction of the choice set. The resulting structural discrete-choice model is consistent with both hedonic and sorting theory.

VI. Extending the Structural Model to Allow a Joint Job-House Choice

All of the structural hedonic/sorting models in the previous section treated wage income as exogenous to location choice. Yet, as chapter 1 emphasized, job opportunities are perhaps the most important determinant of where people live. If households simultaneously sort among communities and labor markets, the supply of public goods will affect behavior in both markets as Rosen (1979) suggested. Thus, one would expect job locations to convey additional identifying information about preferences. This section formally extends the location choice problem to consider a joint job-house choice made by households with heterogeneous job skills and horizontally differentiated preferences for public goods. Part (A) discusses how this approach allows a more flexible depiction of the spatial landscape and brings new information to the problem. Then part (B) formalizes the choice problem and uses a single-crossing restriction to compare the implications for sorting behavior with those in the vertically differentiated model developed by Epple and Sieg (1999).

(A) Rethinking the Spatial Landscape

Using a structural model to explain housing and labor market choices has the advantage of allowing complete flexibility in how the choice set is defined. This extension makes it possible to redefine the spatial landscape in a way that combines the desirable features of the

different models discussed in this chapter. As in the interregional model, the K metropolitan areas are treated as a set of labor markets that differ in the wage paid to workers with each job skill. Assuming J housing communities and K labor markets, each (j,k) pair represents a unique job-house combination which will also be referred to as a “location”. Each location requires a specific commute so that commuting time can be added to the model as an additional location-specific characteristic. Thus, the same community with a different commute time to a different job site represents a different location. For a household that commutes between community j and job location k , let $w_{j,k}(\theta)$ represent wage earnings less the value of time spent commuting. θ indexes both job skill and the shadow value of time. Then, a household’s income equals $\hat{y} + w_{j,k}(\theta)$, its exogenous non-wage income (\hat{y}) plus its “virtual wage income”. Holding their job location and nominal wage fixed, a household can vary its virtual wage income by changing its home community, and therefore its commute time.

A household’s chosen commuting pattern can provide additional information about its preferences for public goods. For example, suppose a household prefers the public goods provided by suburban communities located far from its job. It is forced to choose between a desirable community with a long commute and a less desirable community with a shorter commute. Its choice will help to identify the strength of its preferences. This is just another application of Samuelson’s (1948) revealed preference logic. The primary difference between this case and the one described in the section V.A is that here a household’s choice bounds its *indirect* indifference curve in (y, g) space.

Figure 2.2 contrasts the newly redefined spatial landscape (panel D) with the other models that have been discussed in this chapter. Recall that metro areas 1 and 4 are each individual housing markets, while metro areas 2 and 3 are considered part of the same housing market ($q = 2$). Panels A and B show how the spatial landscape is depicted in the property value and interregional hedonic models, and panel C illustrates the landscape in the traditional sorting model. In panel D, the asterisks show the location of the central business district in each metropolitan area and the arrows illustrate the feasible set of home-to-work commuting patterns. The choice set is confined to a single housing market, as in the property

value hedonic and sorting models. Three “margins of adjustment” help to identify a household’s preferences: (1) choice of home metropolitan area; (2) choice of community within the home metropolitan area; (3) choice of work metropolitan area. These choices uniquely determine the time spent commuting.

Relative to the interregional hedonic model, the depiction of the landscape in figure 2.2D makes it possible to relax the national mobility assumption and still use variation in wages to help identify preferences. Measuring variation in virtual wage income at the intensive margin eliminates the need for data on variation in nominal wage income at the extensive margin. A second difference is that the new landscape recognizes that households may have the opportunity to live in one metropolitan area and work in another, while the interregional model does not.

Previous structural models of location choice have included commuting distance as a locational characteristic while treating work destinations as exogenous (e.g. Bajari and Kahn [2005] and Bayer et al. [2005]). Compared to those models, the treatment of the landscape in panel D differs in that it recognizes households have the ability to change the location of their job²¹. However, this flexibility only brings new information to the problem if the housing market includes more than one possible job destination, like market #2. Most households in the United States live in areas that appear to fit this description. More than 80% of the U.S. population lives in the metropolitan statistical areas (MSA) shown in Figure 2.11. Notice that most of these MSAs are adjacent to at least one other MSA. It seems reasonable to expect that in many cases adjacent MSAs could be considered part of the same housing market.

(B) Redefining the Location Choice Problem

To formalize the location choice problem, let $L_{j,k}$ denote a unique job-house location.

Utility maximization is similar to the sorting model in (2.27)-(2.28), except that households now optimize over two dimensions of location choice and a budget constraint that varies

²¹ So, Orazem and Otto (2001) use a random utility framework to explain job-house choices based on differences in wages, prices, and commutes between metropolitan and non-metropolitan areas, but their analysis does not include local public goods.

across locations. Equation (2.40) describes the utility maximization problem for household i .

$$(2.40) \quad L_{j,k}^* = \max_{j,k} V[g_j(\gamma_i), p_j, \alpha_i, y_{i,j,k}], \quad \text{where } y_{i,j,k} = \hat{y}_i + w_{j,k}(\theta_i).$$

Holding the community fixed at j , a utility-maximizing household will always choose to work in the labor market that provides it with the highest effective wage income, given its job skills. Let $\hat{w}_j(\theta)$ represent the maximum effective wage income that can be obtained by a household living in community j . Then (2.40) can be rewritten as (2.41), with k optimized out of the expression.

$$(2.41) \quad L_j^* = \max_j V[g_j(\gamma_i), p_j, \alpha_i, y_{i,j}], \quad \text{where } y_{i,j} = \hat{y}_i + \hat{w}_j(\theta_i).$$

For each (γ, θ) “type” household, the relevant choice set can be further reduced to a subset of the J communities. This is because, conditional on values for γ and θ , some communities may be dominated. A community is dominated if there is another with more public goods and either a sufficiently lower price, a sufficiently higher effective wage, or both. For example, given $g_1(\gamma) > g_2(\gamma)$, community 1 dominates community 2 if prices and effective wages are defined such that: $p_1 < p_2$ and $\hat{w}_1(\theta) > \hat{w}_2(\theta)$. No utility-maximizing (γ, θ) -type would ever locate in community 2. Let R denote the total number of communities that are not dominated. Then equation (2.42) shows how the relevant choice set for each (γ, θ) -type relates to the set of all communities in (2.41), and to the set of all locations in (2.40).

$$(2.42) \quad \{L_1, \dots, L_R \mid \gamma, \theta\} \subset \{L_1, \dots, L_J \mid \gamma, \theta\} \subset \{L_{1,1}, \dots, L_{J,K} \mid \gamma, \theta\}.$$

As before, imposing a single crossing restriction on preferences makes it possible to characterize how, in equilibrium, households of each (γ, θ) -type must be sorted across the R communities that are not dominated for that type.

Assuming the MWTP for public goods per unit of housing is monotonically increasing in $(\hat{y} \mid \alpha, \gamma, \theta)$ and $(\alpha \mid \hat{y}, \gamma, \theta)$, three conditions will characterize sorting by each

household type in a locational equilibrium: *boundary indifference*, *stratification*, and *non-decreasing bundles*²². Without loss of generality, let the R locations be ordered according to their perceived provision of public goods, $g_1(\gamma) < \dots < g_R(\gamma)$. Boundary indifference requires a household on the “border” between two locations in (α, \hat{y}) space to be exactly indifferent between those locations. Equation (2.43) defines the set of border individuals. It must hold for all $r = 1, \dots, R - 1$.

$$(2.43) \quad \left\{ (\alpha, \hat{y} | \gamma, \theta): V[g_r(\gamma), p_r, \alpha, \hat{y}, \hat{w}_r(\theta)] = V[g_{r+1}(\gamma), p_{r+1}, \alpha, \hat{y}, \hat{w}_{r+1}(\theta)] \right\}.$$

The non-decreasing bundles property requires that for any two locations in the ordering, $(r, r + 1)$ equation (2.44) must hold.

$$(2.44) \quad g_{r+1}(\gamma) > g_r(\gamma) \Rightarrow p_{r+1} > p_r \quad \text{or} \quad \frac{1}{\hat{w}_{r+1}(\theta)} > \frac{1}{\hat{w}_r(\theta)} \quad \text{or both}.$$

The equation implies that households must “pay” for the additional public goods provided by higher ranked locations through housing prices, effective wage income, or both. The third property, stratification, requires that households of each type are stratified across the R ordered locations by $(\alpha | \hat{y})$ and by $(\hat{y} | \alpha)$, as defined in (2.45).

$$(2.45) \quad \begin{aligned} & (\hat{y}_{r-1} | \alpha, \gamma, \theta) < (\hat{y}_r | \alpha, \gamma, \theta) < (\hat{y}_{r+1} | \alpha, \gamma, \theta) \\ & \text{and} \\ & (\alpha_{r-1} | \hat{y}, \gamma, \theta) < (\alpha_r | \hat{y}, \gamma, \theta) < (\alpha_{r+1} | \hat{y}, \gamma, \theta) \end{aligned}.$$

Boundary indifference, *stratification*, and *non-decreasing bundles* describe the partition of preference space that rationalizes equilibrium location choices. Non-decreasing bundles identifies locations that have adjacent regions in the partition. Boundary indifference defines the borders that divide those regions, and stratification guarantees that

²² Boundary indifference and stratification follow from the proof of proposition 1 in Epple and Sieg (1999) because income is separable in non-wage income and effective wage income. To see why non-decreasing bundles must hold, suppose equation (10) fails for some $(r, r+1)$ pair. Then r must have fewer perceived public goods, more expensive housing, and lower effective wage income. If so, $r+1$ dominates r , which implies $r \notin R$, a contradiction.

each region is connected in $(\alpha | \hat{y}, \gamma, \theta)$. In the special case where wage income is exogenous and households are vertically differentiated, the three properties reduce to the ones from the Epple-Sieg (1999) model shown in (2.31)-(2.33).

Translating this theoretical model into an estimable empirical framework presents some new econometric and data-related challenges. Some of these issues include: the need to express wage income as a function of job skill and the opportunity cost of time; job locations and commuting patterns must be observed; and the wage rate each household would earn at each alternative job location must be specified. The next two chapters focus on these and other econometric and data issues.

VII. Conclusion

This chapter reviewed the literature on using households' location choices to infer their demand for public goods. The discussion started by asserting that households get utility from the characteristics of their house and from the public goods provided at that location. This premise forms the basis for hedonic and sorting models that depict a locational equilibrium and then use the properties of that equilibrium to infer households' preferences. The various econometric strategies for estimating each class of model were described in a single notational framework, with special attention to the restrictions they place on the structure of preferences. While hedonic and sorting models differ in how they depict the supply side of a locational equilibrium, they share a common representation of consumer behavior. This observation was used to argue that the difference between a "structural hedonic model" and a "structural sorting model" is purely semantic since both estimators treat supply as exogenous and rely entirely on consumer behavior to identify preferences. Finally, the location choice problem was extended to recognize that households choose both job and house locations. Equilibrium sorting behavior was described under the assumption that households differ in their job skills and in their preferences for multiple public goods.

Every model in this chapter places restrictions on the structure of preferences. For the most part, analysis must treat these as maintained assumptions. Without imposing some

types of structural restrictions, the most that can be learned about a household's demand for a public good is their marginal willingness-to-pay. To recover the willingness-to-pay for a non-marginal change, the analyst must introduce additional information. Advances in microeconomic methods enhance the flexibility in how this information can be provided. There is a clear tradeoff between restrictions and data needs. This can be illustrated by comparing the alternative formulations of the hedonic model. At one extreme, the "multiple-market" approach only imposes a general restriction on preferences. Estimation in this case requires data from multiple markets and fairly extensive restrictions on the data generating process. At the other extreme, Bajari and Benkard (2005) are able to identify preferences using minimal data from a single market but they must fully specify the utility function. The strategy suggested by Ekeland et al. (2004) falls somewhere in the middle.

If the analyst chooses to specify the utility function the resulting structural estimator could accurately be labeled a "sorting" model or a "hedonic" model. The three structural estimators discussed in the chapter (Epple-Sieg [1999]; Bajari-Benkard [2005]; Bayer et al. [2005]) are distinguished only by the restrictions they impose on the indirect utility function and the assumptions they place on the distribution of each heterogeneous preference parameter. Chapter 3 extends this class of estimators, recognizing that spatial variation in local public goods may help to determine job locations as well as house locations. This provides a structural analog to the interregional hedonic model on a smaller geographic scale.

Table 2.1: Summary of Notation

The Spatial Landscape	
$q = 1, \dots, Q$	Index of housing markets
$k = 1, \dots, K$	Index of metropolitan areas / labor markets
Locational Characteristics: indexed by the j subscript	
$j = 1, \dots, J$	Index of locations (depending on the model, j = house, community, or metro area)
h	Vector of housing characteristics (e.g. bedrooms, bathrooms, pool)
g	Vector of public goods (e.g. air quality, school quality, open space)
$x \subset [g, h]$	Subset of housing characteristics and public goods observed by the econometrician
ξ	Vector of characteristics unobserved by the econometrician
$\bar{h} = f(h)$	Index of housing characteristics
$\bar{g} = f(g)$	Index of public goods
$\bar{\xi} = f(\xi)$	Index of unobserved characteristics
P	Annualized expenditures on a single home
$\hat{P}_1 = f(x, \xi)$	Marginal price function for x_1 estimated from a hedonic model
p	Annualized “per-unit” price for a homogeneous unit of housing (i.e. the rental rate)
b	Composite numeraire private good. Its price is assumed to be normalized to unity
Household Characteristics: indexed by the i subscript	
$i = 1, \dots, I$	Index of households
$y = \hat{y} + w$	Total annual income
\hat{y}	Annual exogenous non-wage income
w	Annual wage income
$\alpha = f(d, \varepsilon)$	Preferences for locational characteristics
d	Vector of household demographics (e.g. education, age, race, income)
ε	Idiosyncratic preferences for locational characteristics
Reduced Form Estimation	
B	Parameter vector estimated from the hedonic price function
Δ	Parameter vector estimated from the hedonic wage function
z	Instruments for x
Ω	Parameter vector estimated from the inverse demand function

Table 2.2: Restrictions Used to Recover the Demand for a Public Good from a Hedonic Model (“-----” indicates no restriction)

		(1)	(2)	(3)	(4)	(5)
Category	Argument	Multiple Markets	Restricting Preferences	Fully Specified (continuous)	Fully Specified (discrete)	Interregional
Maintained Assumptions	location	House	House	House	House	Metropolitan area
	<u>Choice set:</u> limit on mobility	Housing market	Housing market	Housing market	Housing market	United States
	<u>Choice set:</u> continuity	Continuous	Continuous	Continuous	Discrete	Continuous
Structural Restrictions	Utility function	-----	Marginal utility is monotonically separable in: x, d, ε	Fully specified with $\dim(\alpha) \leq \dim(x)$	(i) fully specified (ii) differentiable (iii) nonsatiated	-----
	Distribution of preferences $f(\alpha)$	$f_1(d) = \dots = f_Q(d)$	-----	-----	$f(\alpha)$ specified	$\alpha_1 = \dots = \alpha_I$
Econometric Restrictions	Unobserved location attributes (ξ)	$E[\xi x] = 0$	$\xi = 0$	$\xi \perp x$	$\xi \perp x$	$E[\xi x] = 0$
	Instruments for location attributes (z)	$E[z'x]$ is full rank	-----	-----	-----	-----
	Idiosyncratic preferences (ε)	$E[\varepsilon z, d] = 0$	$E[\varepsilon z, d] = 0$	-----	-----	-----
Key Conceptual Work		Rosen (1974) Brown-Rosen (1982) Epple (1987)	Ekeland et al. (2004)	Bajari-Benkard (2005)	Bajari-Benkard (2005)	Roback (1982) Blomquist et al (1988)
Location Choice Application		Palmquist (1984)	None	Bajari-Kahn (2005)	None	Blomquist et al (1988)

Table 2.3: An Example of Housing Prices and Air Quality in Four Communities

House Location	Air quality	Price
1	1.25	1.00
2	1.85	1.25
3	1.66	1.26
4	2.00	1.50

* Higher values indicate higher quality.

Table 2.4: An Example of Prices and Two Public Goods in Four Communities

Community	Public Goods*		Price
	Air quality	School quality	
1	1.25	1.25	1.00
2	1.85	1.65	1.25
3	1.66	1.86	1.26
4	2.00	2.00	1.50

* Higher values indicate higher quality.

Table 2.5: Adding a Fifth Community to the Choice Set

Community	Public Goods*		Price
	Air quality	School quality	
1	1.25	1.25	1.00
2	1.85	1.65	1.25
3	1.66	1.86	1.26
4	2.00	2.00	1.50
5	1.95	1.90	1.64

* Higher values indicate higher quality.

Table 2.6: Restrictions Used to Recover Preferences in Structural Hedonic and Sorting Models

Category	Argument	Epple-Sieg (1999)	Bajari-Benkard (2005)	Bayer et al. (2005)
Maintained Assumptions	Income	Exogenous to location choice	Exogenous to location choice	Exogenous to location choice
	Free mobility	yes	yes	yes
	Set of public goods	Discrete	Discrete	Discrete
	Set of housing characteristics	Continuous	Continuous / Discrete	Discrete
Restrictions on Preference Heterogeneity	Preferences for observed characteristics	Vertical	Horizontal	Horizontal
	Preferences for unobserved characteristics	Vertical	Vertical	Horizontal
	Distributional assumptions	$F(\alpha, y) \sim \text{lognormal}$	None	$\varepsilon \sim \text{iid logit}$
Restrictions on Substitution Possibilities	Public goods w/ public goods	Perfect Substitutes	Perfect Substitutes	Perfect Substitutes
	Public goods w/ housing	CES utility \rightarrow anywhere from perfect substitutes to perfect complements	Perfect Substitutes	Perfect Substitutes
	Substitutes per location (J total locations)	2	Between 2 and $J-1$	$J-1$
Source of Identification for Unobserved Public Goods ($\bar{\xi}_j$)		$F(\alpha, y) \sim \text{lognormal}$	Hedonic price function	$\varepsilon \sim \text{iid logit}$

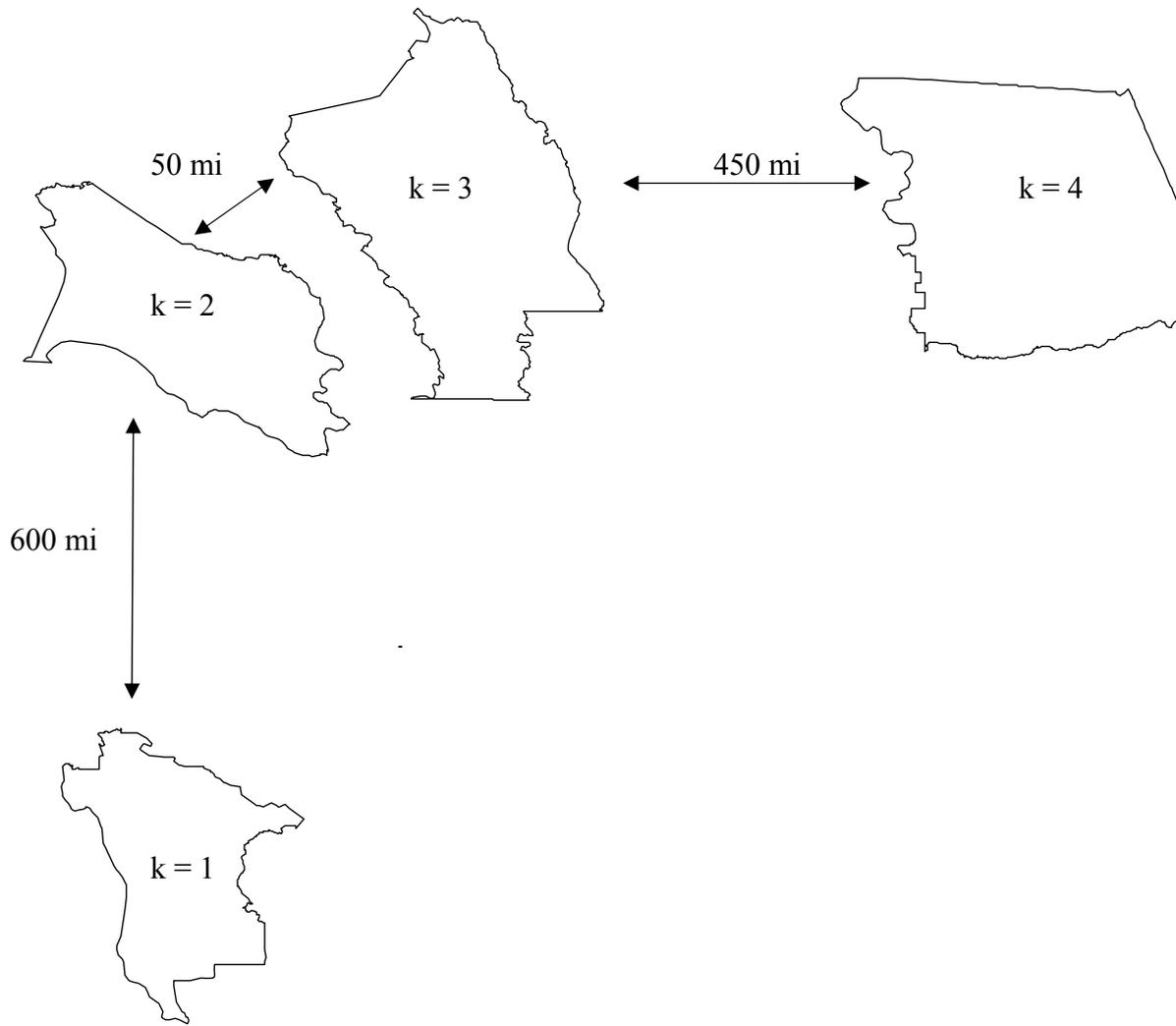


Figure 2.1: A Stylized Depiction of the Urban Landscape: 4 Metropolitan Areas

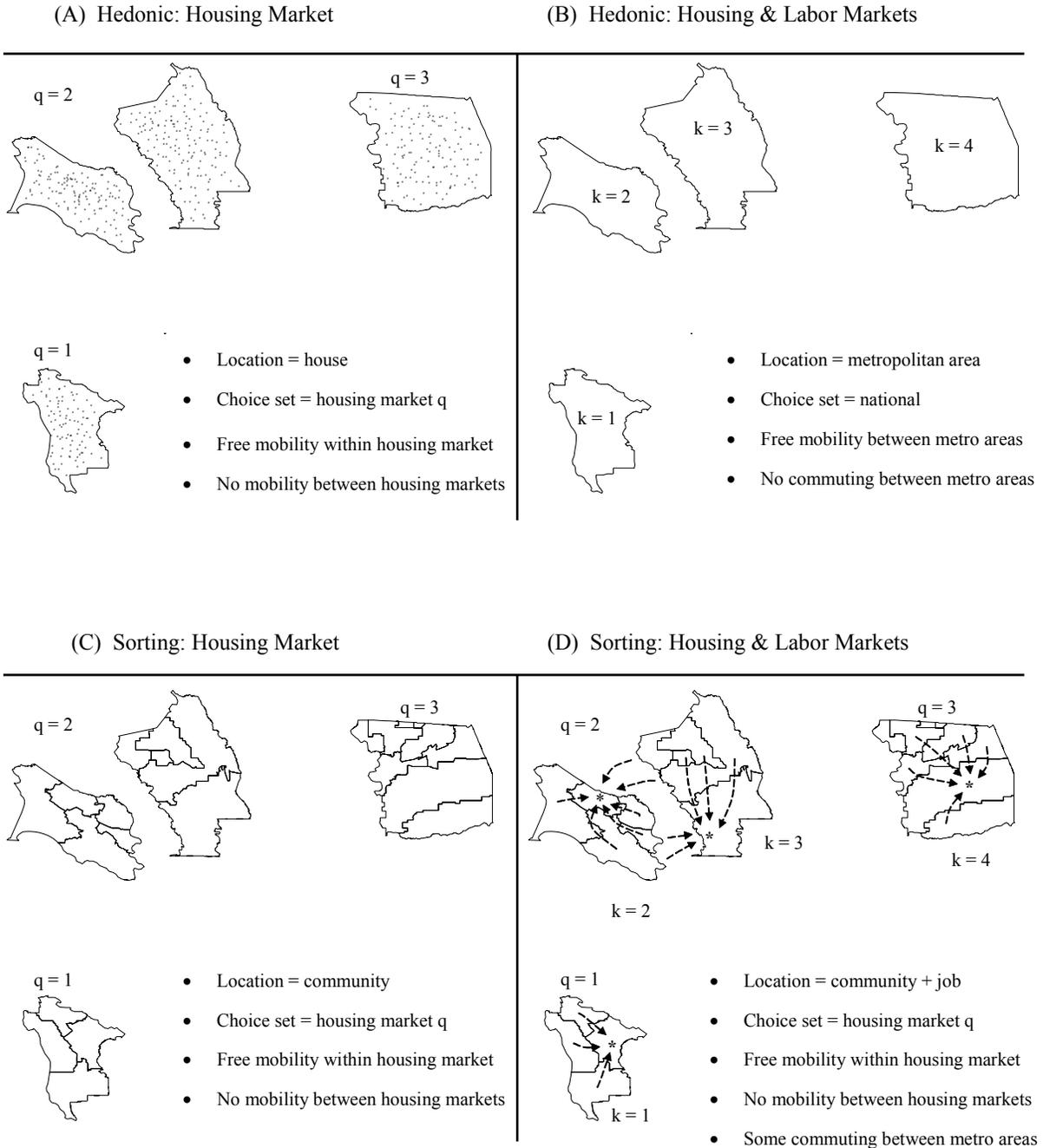


Figure 2.2: Defining Locations, Markets, and Mobility in Hedonic and Sorting Models

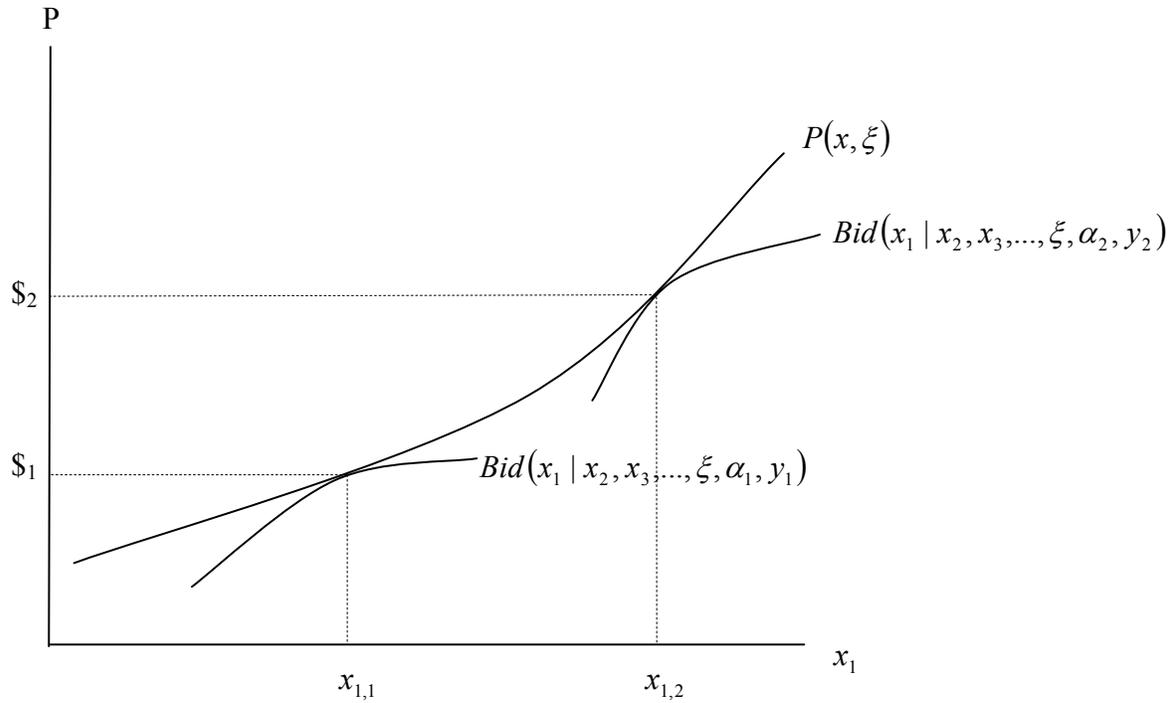


Figure 2.3: Bid Functions for Housing as a Function of x_1 in Hedonic Equilibrium

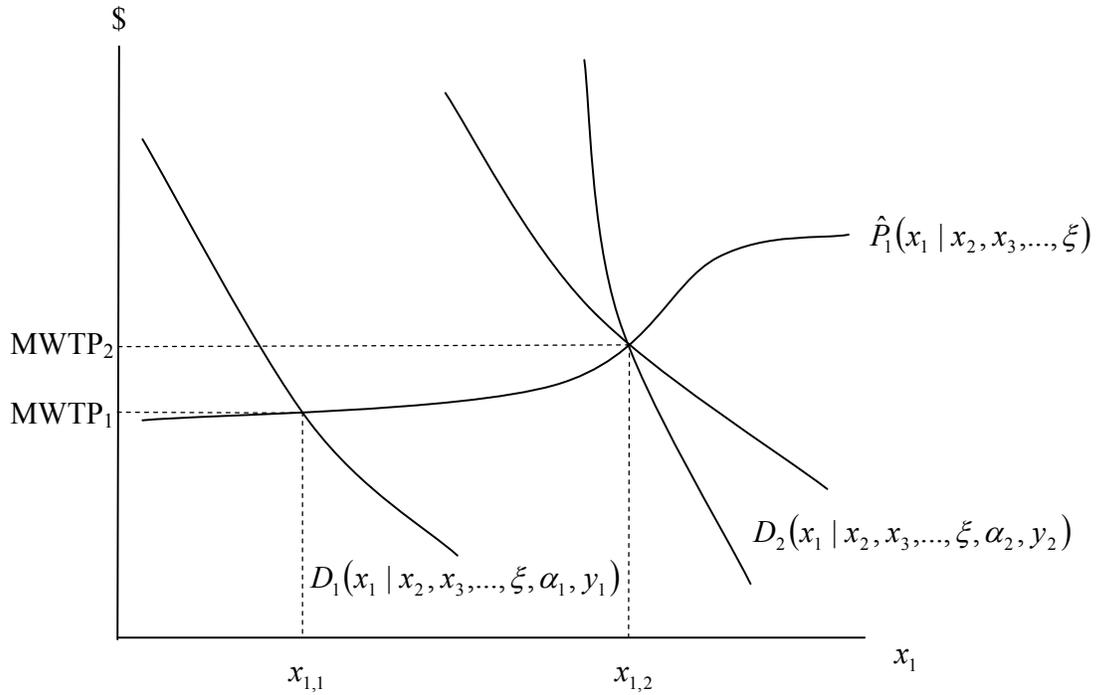


Figure 2.4: Implicit Price Function for x_1 and Demand Curves for Two Households

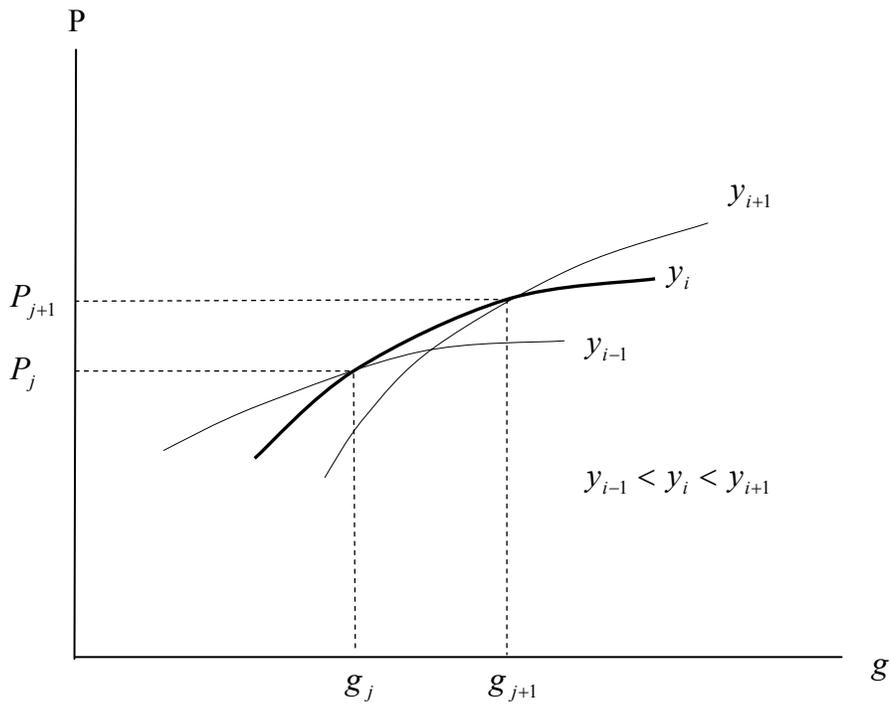


Figure 2.5: The Single Crossing Condition: Indifference Curves for Three Households



Figure 2.6: Partition of Households across Communities by Income

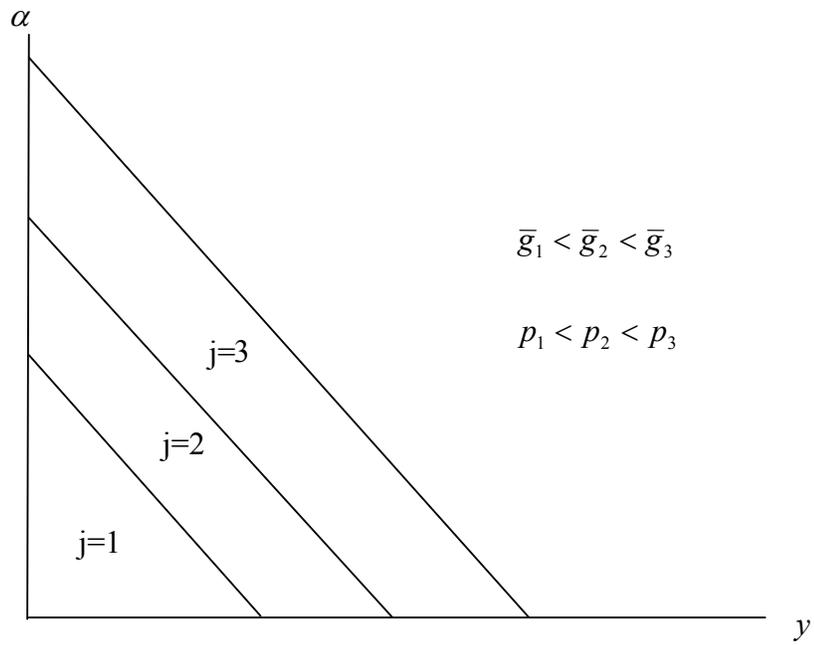


Figure 2.7: Partition of Households into Communities in Epple-Sieg (1999)

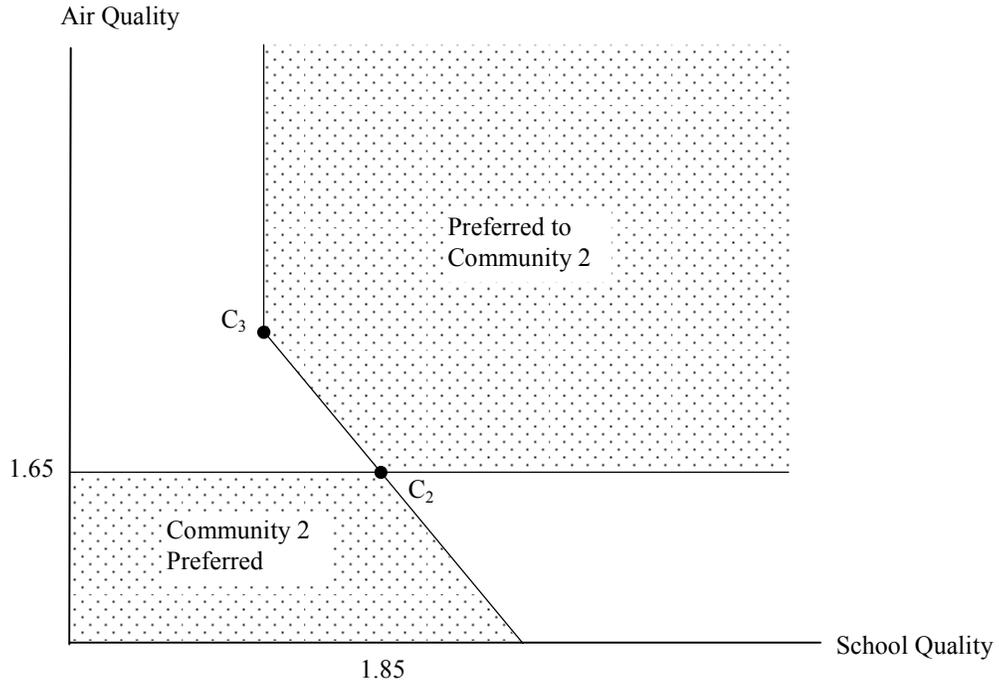


Figure 2.8: Revealed Preference Logic in Characteristics Space

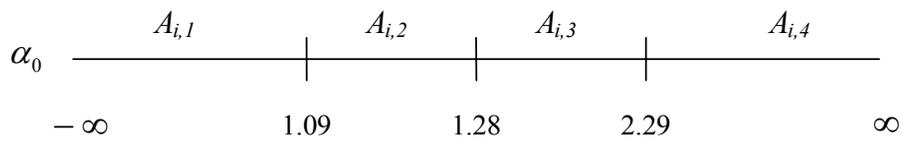


Figure 2.9: Partition of α at $y=\$50,000$

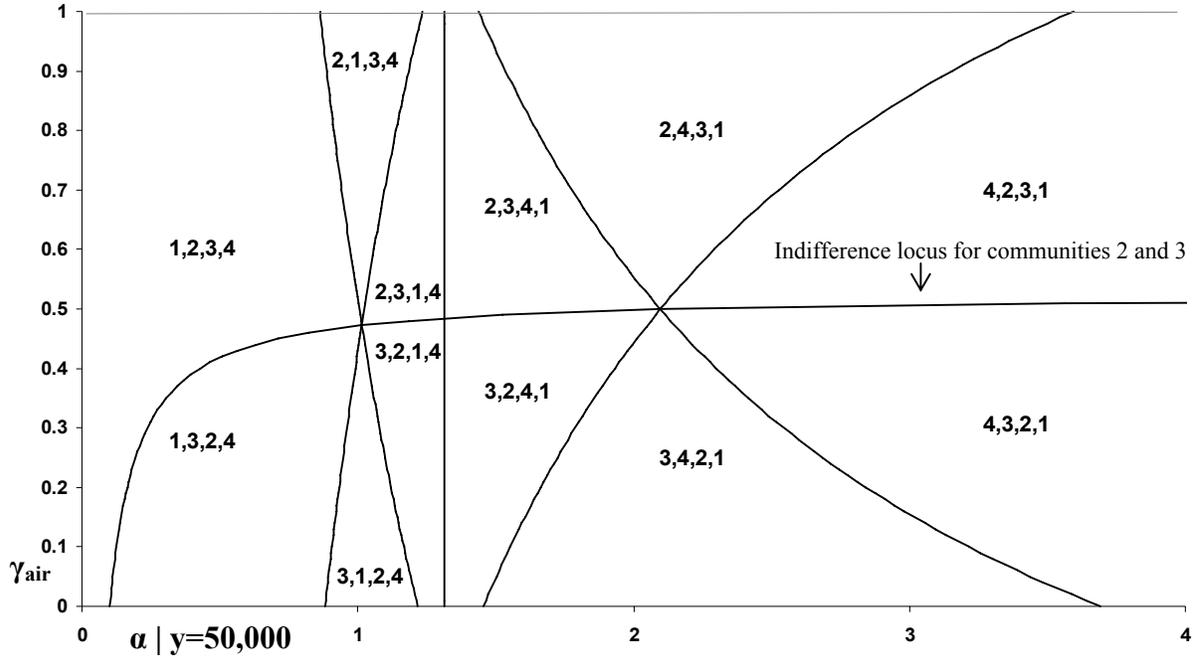


Figure 2.10A: Boundary Indifference Loci and Community Rankings by Utility
 $(\gamma_{\text{school}} + \gamma_{\text{air}} = 1)$

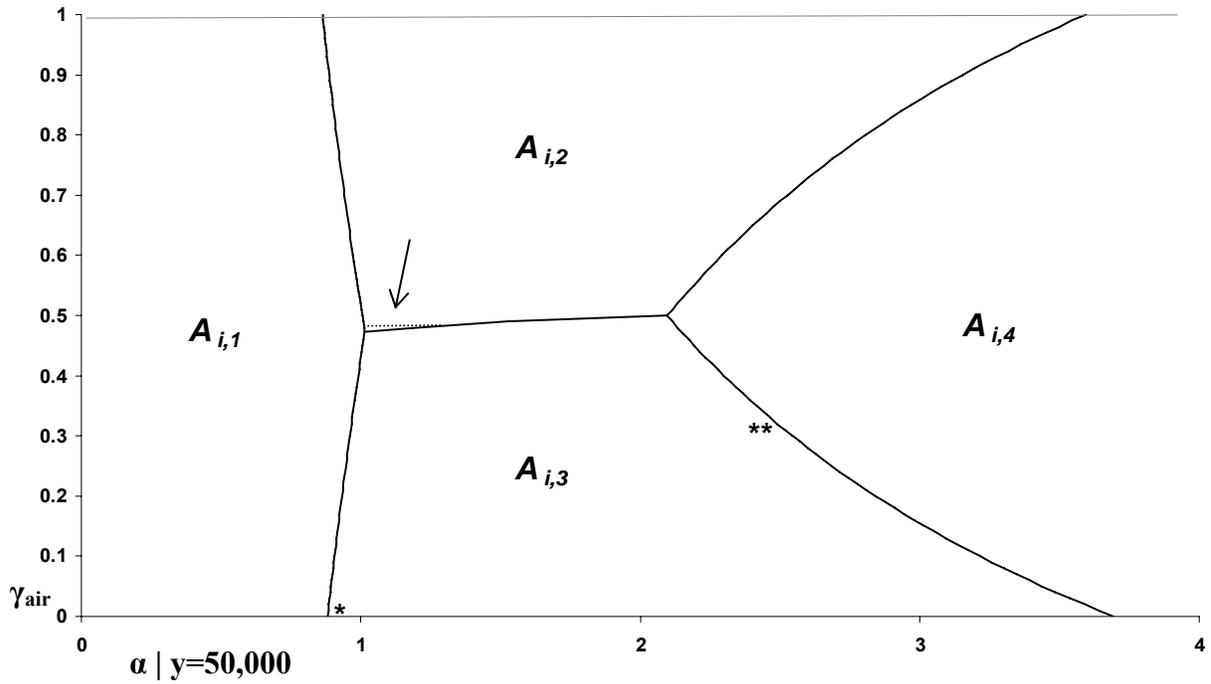


Figure 2.10B: Partitioning Preference Space: Horizontally Differentiated Preferences

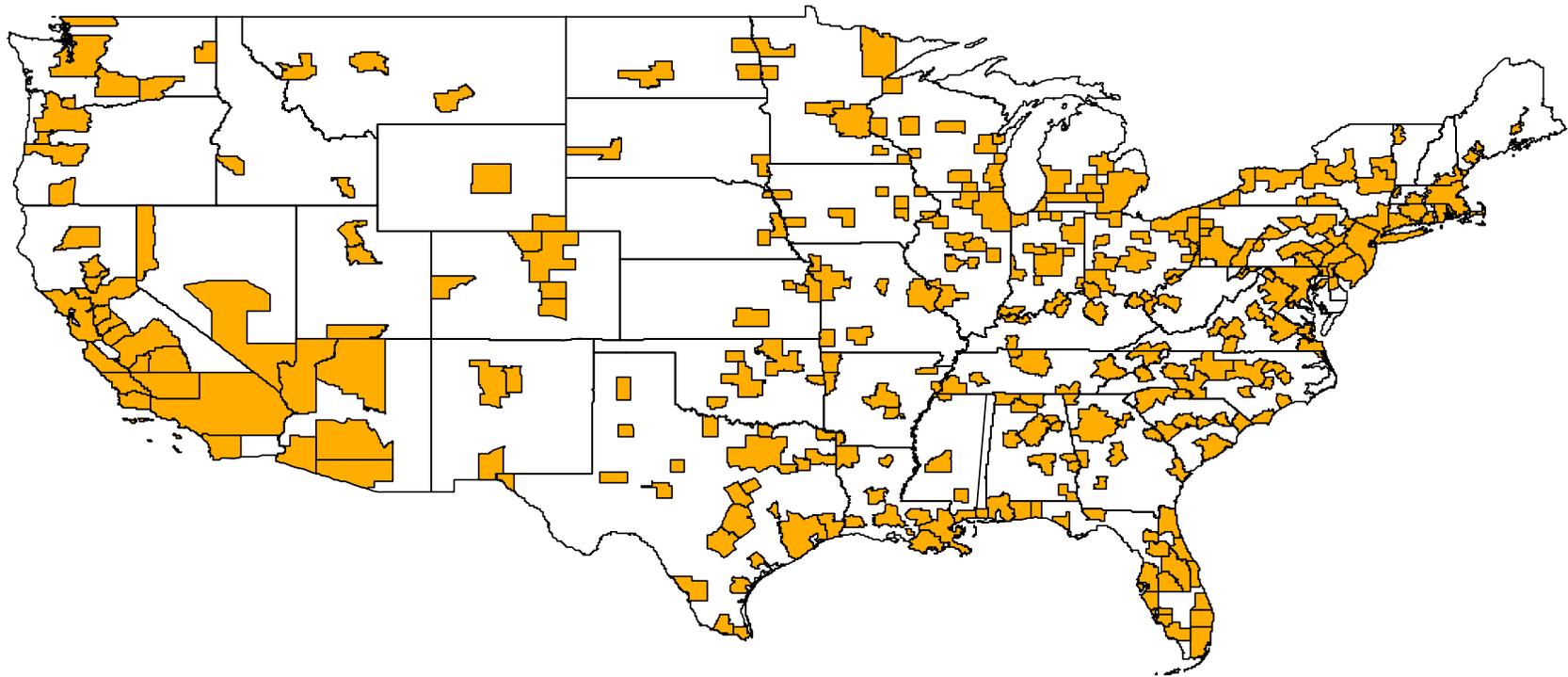


Figure 2.11: Metropolitan Statistical Areas in the United States

Chapter 3: Empirical Model

I. Introduction

This chapter develops an empirical sorting model to explain how local public goods influence where households choose to live and work. By including work destinations as a dimension of location choice, the new “dual-market” framework extends the sorting models described in the previous chapter. The existing empirical literature uses the location of a household’s home to infer its preferences for the public goods provided at that location. This approach treats job locations and wage income as fixed. Chapter 2 extended the underlying theoretical model to recognize that, for working households, job and house locations are connected because a worker’s wage-commute options depend on the location of his home. Observing the household’s location choices in both markets offers the potential to learn more about its preferences for public goods than observing its house location alone. The goal of this chapter is to translate the dual-market sorting logic into an empirical model that can be estimated to recover households’ preferences for local public goods.

In practice, the analyst can observe some of the characteristics of a household (such as income, occupation, and demographics) and some of the characteristics of its location choices (such as housing prices, public goods, and wages). In order to use this information to infer the household’s preferences for public goods, the analyst must first define the set of choice alternatives and a specification for the utility function. Once defined, these are treated as maintained assumptions on the econometric model. Sections II and III of this chapter discuss how these maintained assumptions influence what can be learned about the household’s preferences and their willingness-to-pay for improvements in public goods.

Section II begins by proposing a specification for the indirect utility function. The key features of that specification are contrasted with the empirical sorting models in chapter 2. The most important departure from the existing literature is the treatment of wage income. The specification allows different working households to face different wage-commute options, depending on their job skill. The work destinations they choose provide additional information about their preferences for local public goods.

Given the specification for indirect utility, the definition for the set of choice alternatives determines what location choices can reveal about preferences. The analyst must

provide definitions for a “housing community”, a “work destination”, and geographic boundaries for the study region. Section III of the chapter discusses how the underlying sorting theory implies some guidelines for how these concepts are defined.

After defining the choice set and the indirect utility function, the challenge is to develop an econometric strategy that uses all the available data on households and their location choices to recover the parameters of the indirect utility function. This task is the focus of sections IV and V. Section IV outlines a two-stage strategy to recover all the parameters of the model, using subsections to explain how each parameter is identified. Then section V elaborates on the computational details. Finally, section VI concludes by summarizing the key features of the new dual-market framework and comparing it with alternative models. Chapters 4, 5, and 6 use the new framework to analyze the demand for air quality in Northern California.

II. Indirect Utility Function

The first step in developing the empirical model is to choose a specification for the indirect utility function that captures the important features of the dual-market sorting problem. There are at least three important modeling issues to consider: the depiction of preference heterogeneity, the treatment of labor market choices, and the relationship between public and private goods. As noted in chapter 2, the depiction of preference heterogeneity controls the scope of substitution patterns, affects how the welfare measures can be interpreted, and determines what can be learned about relative preferences for different public goods. Generalizing the model to allow labor market choices provides more information to identify preferences. It also supports more complex patterns of sorting behavior and introduces another extensive margin along which households may adjust to large-scale shocks. Adjustment along the intensive margin will depend on households’ willingness to trade public goods for private goods.

The urban landscape is assumed to consist of $j = 1, \dots, J$ communities, each of which is distinguished by its provision of local public goods and the price of housing. The

landscape is also divided into $k = 1, \dots, K$ labor markets, each of which offers a different wage to workers with each job skill. Each working household chooses a particular $(j, k) = (\text{house}, \text{job})$ location. To simplify notation in what follows, let locations $j, k = (1, 1), \dots, (J, K)$ be indexed by $z = 1, \dots, Z$, so that z represents a unique (job, house) combination.

Working households are assumed to choose work destinations on the basis of job opportunities available to their primary earner. Wage income for all other members of the household is treated as exogenous²³. Each worker possesses one of S different observable occupations and every household may differ in its preferences $(\alpha_i, \gamma_i, \theta_i)$, so households are indexed by both i and s . Then the indirect utility obtained by household i, s in location z can be expressed as (3.1).

$$(3.1) \quad V_{i,s,z} = \left\{ \alpha_i (\bar{g}_{i,z})^\rho + \left[\exp\left(\frac{y_{i,s,z}^{1-\nu} - 1}{1-\nu}\right) \exp\left(-\frac{\beta p_z^{\eta+1} - 1}{1+\eta}\right) \right]^\rho \right\}^{\frac{1}{\rho}},$$

where $\bar{g}_{i,z} = \gamma_{i,1} g_{1,z} + \dots + \gamma_{i,N-1} g_{N-1,z} + \gamma_{i,N} \bar{g}_z$, and

$$y_{i,s,z} = \hat{y}_i + \theta_{i,1} w_{s,z} (1 - \theta_{i,2} t_{s,z}).$$

This CES function generalizes the specification from Epple and Sieg (1999) to allow households to differ in their relative preferences for multiple public goods and to allow wage income to vary across job locations. Parts (A) and (B) of this section highlight the key features of each term, and part (C) discusses the relationship between the two terms. Finally, part (D) discusses the implications for allowing a working household's income to vary with their work destination in a way that depends on their job skill ($\theta_{i,1}$) and the time spent commuting ($t_{s,z}$). Much of the discussion focuses on how (3.1) relates to the alternative specifications used by Epple and Sieg (1999), Bajari and Benkard (2005), and Bayer, McMillan, and Reuben (2005), henceforth ES, BB, and BMR.

²³ Chapter 4 discusses the consequences of this simplification, and chapter 7 returns to this issue by considering a dual-earner job search as a topic for future research.

(A) *Utility from Public Goods*

Households differ in their preferences for a bundle of N public goods provided by each community. $N-1$ of the public goods are observable. The N^{th} public good ($g_{N,z} = \bar{\xi}_z$) represents an index of all the community-specific attributes that are not observed by the econometrician. Households are restricted to have the same objective evaluation for this index; that is, the weights in the index are assumed to be the same for all households ($\omega = \omega_i \forall i$). This allows $\bar{\xi}_z$ to be treated as if it were a single public good.

Equation (3.2) shows how the bundle of public goods enters the utility function through the linear index, $\bar{g}_{i,z}$. As in chapter 2, the bar superscripts indicate that $\bar{g}_{i,z}$ and $\bar{\xi}_z$ are indices.

$$(3.2) \quad V_{i,s,z} = \left\{ \alpha_i (\bar{g}_{i,z})^\rho + [\cdot]^\rho \right\}^{\frac{1}{\rho}},$$

$$\text{where } \bar{g}_{i,z} = \gamma_{i,1}g_{1,z} + \dots + \gamma_{i,N-1}g_{N-1,z} + \gamma_{i,N}\bar{\xi}_z, \quad \text{and } \bar{\xi}_z = f(\omega, \xi_z).$$

Households differ in the weights they place on each public good in the index ($\gamma_{i,1}, \dots, \gamma_{i,N}$) and in their overall preferences for public goods relative to private goods (α_i). α_i is not identified separately from $\gamma_{i,1}, \dots, \gamma_{i,N}$. To distinguish these effects the weights are assumed to sum to 1, allowing α_i to be identified separately as a scaling parameter on the strength of preferences. While α_i is extraneous, there are two reasons for including it in the specification. First, normalizing the weights in the public good index simplifies the computational model²⁴. Second, it allows a direct comparison to the depiction of preference heterogeneity in ES.

²⁴ Normalizing the weights to sum to 1 allows the value for each of the weights to be bounded by 0 and 1 for every household. This greatly simplifies the series of rootfinding problems used to recover the γ 's numerically during the Gibbs sampling algorithm described in sections IV and V. Since α can be recovered analytically, allowing it to be unbounded does not pose any inconvenience.

Relative to other sorting models, the specification takes an intermediate stance in its depiction of preference heterogeneity. Households have horizontally differentiated preferences for the observed public goods and vertically differentiated preferences for the unobserved public goods (as in BB). This implies the number of substitutes for each location ranges from 2 and Z , compared to exactly 2 in ES and exactly Z in BMR. The public goods index in (3.2) simplifies to the vertically differentiated specification used by ES in the special case where households have identical relative preferences for the different public goods; i.e. when the i subscript is dropped from $(\gamma_1, \gamma_2, \dots, \gamma_N)$. In contrast, generalizing the public goods index to reproduce the specification used by BMR would require adding idiosyncratic location-specific tastes for each household; i.e. adding i subscripts to the ω 's.

Another feature that differentiates the depiction of preference heterogeneity from ES and BMR is the absence of maintained assumptions on the distribution of preferences. Recall that the econometric strategies developed by ES and BMR rely on a priori parametric assumptions on the distributions of heterogeneous preference parameters. The observation that those parameters are set identified by the partitioning logic described in chapter 2 suggests an alternative strategy for estimation: first recover the region of preference space that justifies each location choice, then sample from that region according to the desired assumption about the distribution of preferences. The econometric model in section IV explains how to implement this strategy.

(B) *Utility from Private Good Component of Housing*

The second term in the utility function simplifies to the specification in ES in the special case where income does not vary with the household's location. Both specifications assume that all households share the same values for the housing demand parameters: price elasticity (η), income elasticity (ν), and demand intercept (β). Applying Roy's Identity to the indirect utility function yields the demand function for the private good component of housing (\bar{h}) in (3.3).

$$(3.3) \quad \bar{h}_i = \beta p_z^\eta y_i^\nu, \quad \text{with } \beta > 0.$$

Recall that ES-type sorting models treat housing as a homogeneous commodity that can be consumed in continuous quantities. Under this assumption, the price of housing reflects the cost of consuming the bundle of public goods provided by each community. That is, $p_z = f(g_z, \xi_z)$. However, in practice, observed housing prices will also depend on the structural characteristics of the house. As a result, constructing a consistent measure of p_z requires an additional restriction on preferences. Sieg, Smith, Banzhaf, and Walsh (2002) demonstrate that this requirement is satisfied if the structural characteristics of housing enter the direct utility function through a sub-function that is homogeneous of degree one and separable from the effect of public goods and the numeraire. This restriction is treated as an additional maintained assumption.

While other sorting frameworks define housing as a heterogeneous commodity, they use linear utility functions that impose the same homogeneity and separability restrictions on its structural characteristics (e.g. Bajari and Kahn [2005]; BMR). The linear utility functions in these models also require households to view public and private goods as perfect substitutes, whereas the CES specification in (3.1) relaxes the perfect substitutability restriction.

(C) *Substitution between Public and Private Goods*

All households are assumed to share the same (constant) elasticity of substitution between public and private goods (σ):

$$(3.4) \quad \sigma = \frac{1}{1 - \rho}.$$

While the estimator does not restrict the sign of the CES parameter (ρ) in equation (3.1), the single crossing restriction used to characterize sorting behavior in the theoretical model implies $\rho < 0$ given the expected signs for the housing demand parameters. Restricting ρ to be negative limits the substitutability between public and private goods. As the definition

for σ in equation (3.4) implies, σ is restricted to be less than 1. The practical implication of this restriction is that it guarantees the MWTP for each public good, per unit of housing, will be strictly increasing in income given preferences^{25,26}. Alternatively, if the estimated value for ρ exceeds 0 the MWTP for public goods could be decreasing in income for some values of the heterogeneous preference parameters. Thus, the estimated sign of ρ can be interpreted as a consistency check on the model.

In the existing applications of Epple and Sieg's (1999) model, the elasticity of substitution represents the only link between public and private goods. Otherwise, the utility function is separable in its public and private components. The utility function in (3.1) starts to relax this separability by allowing the demand for housing in community j to depend on its proximity to work destinations. Intuitively, longer commute times decrease the demand for housing by decreasing income.

(D) *Income*

Income is defined by the sum of exogenous non-wage income (\hat{y}) and wage income, less the value of time spent commuting. Equation (3.5) rewrites the expression for income from (3.1) to separate the three terms.

$$(3.5) \quad y_{i,s,z} = \hat{y}_i + \theta_{i,1} w_{s,z} - \theta_{i,2} t_{s,z} (\theta_{i,1} w_{s,z}).$$

The second term in the expression, wage income, depends on the household's job skills. The primary earner of each household is assumed to possess skills that qualify them for an occupation (e.g. biomedical engineer, locksmith, etc). This is the observable component of job skill indexed by s . In the labor market represented by z , the average wage for that occupation is $w_{s,z}$. However, a worker's ability to collect that wage if they were to move from their current job depends on unobservable components of their job skill (e.g. quality of

²⁵ This follows from the discussion of the single-crossing condition in chapter 2 (see equation [2.30]).

²⁶ In contrast, the construction of the indirect utility function in (3.1) requires the MWTP for each public good, per unit of housing, to be strictly increasing in α given income and the other heterogeneous parameters.

education, experience, “people skills”, etc.) and on unobservable attributes of the job. All the unobservables are reflected in a single parameter, $\theta_{i,1}$, that represents the worker’s labor market mobility. If $\theta_{i,1}$ is greater (less) than 1, the worker earns more (less) than the average wage for their occupation.

The third term in (3.5) represents commuting costs. $t_{s,z}$ is the ratio of commute time to work time, and $\theta_{i,2}$ represents the household’s shadow value of time as a share of their wage rate. If $\theta_{i,2} = 0$, commuting is costless. This value for the parameter could represent a case where the worker’s firm pays for their transportation, for example. At the other extreme, if $\theta_{i,2} = 1$, the worker’s shadow value of time equals their wage rate. To the extent that wages and commute times vary across space, income will be endogenous to location choice.

Unobserved heterogeneity in job skill is the most important feature of the income specification. It determines the wage a worker would earn if they were to change jobs and it can help to explain location choices that would otherwise appear irrational. Equation (3.6) illustrates how the job skill parameter determines the wage that would be earned at each location.

$$(3.6) \quad \begin{aligned} y_{i,s,z} &= \hat{y}_i + w_{s,z}(1 - \theta_{i,2}t_{s,z}), & \text{if the worker currently occupies location } z. \\ y_{i,s,z} &= \hat{y}_i + \theta_{i,1}w_{s,z}(1 - \theta_{i,2}t_{s,z}), & \text{otherwise.} \end{aligned}$$

At the worker’s current job $\theta_{i,1} \equiv 1$ because they are observed earning $w_{s,z}$. However, if that worker were to move to a different labor market they may be paid more or less than other workers with the same occupation depending on their relative job skill *within* occupation s .

To see how $\theta_{i,1}$ helps to explain location choices that would otherwise appear irrational, consider the following example. In 2003 the average salary paid to lawyers in the San Jose metropolitan area was \$154,000 compared to \$104,000 in the nearby Santa Cruz metropolitan area (CA Employment Development Dept.). Yet the Census reports 30 lawyers

living in San Jose and commuting to work in Santa Cruz. While this location choice appears irrational, unobserved job skill can provide an explanation: the 30 lawyers would not be eligible to earn \$154,000 if they were to work in San Jose; that is, $\theta_{i,1} < 1$. One interpretation of this result is that earning the market wage in San Jose (or Santa Cruz) requires specialization in a particular field of law. Lawyers who have not specialized in that field may be paid less. In general, this type of wage differential creates an incentive for workers to sort into labor markets according to their (unobserved) skills.

Assuming workers sort into labor markets based on skill, their location choices will help to identify the extent to which the unobserved component of their job skill would allow them to earn the market wage in alternative labor markets. Likewise, their chosen commuting pattern will help to identify their opportunity cost of time spent commuting. Adding these two heterogeneous parameters to the model adds two dimensions to the partition of preference space. While $\theta_{i,1}$ and $\theta_{i,2}$ are both set identified by the same revealed preference logic as α_i and γ_i , the dichotomous treatment of $\theta_{i,1}$ in (3.6) means that observed location choices can not identify a lower bound on $\theta_{i,1}$. A worker's job location choice can always be attributed to the trivial explanation that they are not qualified to work anywhere else ($\theta_{i,1} = 0$).

By exploiting unobserved heterogeneity to help explain observed location choices, $\theta_{i,1}$ plays a role that is similar, but not identical to that of the idiosyncratic tastes for locations ($\varepsilon_{i,j}$) in BMR. There are three important differences in its use here. First, the number of household-specific idiosyncratic taste parameters in BMR equals the number of locations, while in the current model each household has a single idiosyncratic skill parameter. Second, the distribution of the ε 's is specified by the researcher while the distribution of $\theta_{i,1}$ can potentially be informed by data on variation in wages within each occupation. Finally, the ε 's are unbounded due to the type I extreme value assumption, while $\theta_{i,1}$ has natural bounds. Assuming wages are nonnegative, $\theta_{i,1}$ must be nonnegative. Likewise, it is bounded from above by the logic of revealed preferences.

Finally, while the treatment of income departs from the existing sorting models, it is not unlike the treatment in reduced-form hedonic wage models. Like the current specification, hedonic wage models recognize that wages vary with attributes of the worker and with attributes of the job (Rosen [1979]). The difference is that hedonic models typically explain variation in wages using data on the observable characteristics of workers (e.g. race, sex, age, education) and on the observable characteristics of their jobs (e.g. non-wage benefits, mortality risks, morbidity risks) whereas the current model compresses most of these characteristics into $\theta_{i,1}$. If data were available for the characteristics of workers and their jobs, the current specification for income could be replaced by a hedonic wage equation. The idea would be to estimate a reduced-form expression for wages and then use it to predict the wage each working household would be paid in each alternative labor market, as in Bayer, Keohane, and Timmins (2006). However, there is no guarantee that the predicted wage rates would be capable of rationalizing observed job locations unless idiosyncratic job-specific “taste” or “skill” parameters were also added to the utility function. Adopting the hedonic approach would also change the set of labor market choices specified for each working household.

III. Defining the Choice Set

Given its income, job skill, and preferences, each household will choose the location that provides it with the highest level of utility. Chapter 2 demonstrated how this choice can provide set identification of the household’s preferences for the public goods that differentiate locations. Part (A) of this section discusses the consequences of expanding or contracting the study region, given a definition for “housing communities” and “labor markets”. Then, given a definition for the study region, part (B) discusses the importance of how housing communities and labor markets are defined.

(A) *Extensive Margin: Geographic Borders of the Study Region*

For location choices to reveal preferences, households must be able to move freely within the study region. The literature on migration decomposes moving costs into three components: the *physical* costs of moving possessions, the *information* costs associated with learning about the attributes and job opportunities in a new location, and the *psychic* costs of moving away from family, friends, and a familiar environment (Herzog Jr. and Schlottmann [1981]). Sorting models treat all three components as being negligible within the study region and very large outside the region. That is, it is assumed to be so expensive to move out of the region that outside locations can be excluded from the choice set. With this logic in mind, the study region should be sufficiently small that the three components of moving costs can be ignored and the region should be isolated from outside alternatives in a way that substantially increases the cost of leaving. For example, the region could be a metropolitan area surrounded by farmland as in the landscape metaphor Starrett (1981) used to motivate his concept of “internal” capitalization. However, as chapter 2 illustrated, relatively few metropolitan areas fit this description; most are located near other metropolitan areas. If the region of interest is not isolated from outside alternatives in an obvious way, the analyst must find other criteria to define its boundaries. This process poses a tradeoff. Expanding the study region provides more information to identify preferences but requires greater reliance on the free mobility assumption.

To see how expanding the geographic borders of the choice set provides more information about preferences, consider the example in figure 3.1. It depicts a study region that consists of three communities (A, B, and C) which differ in their provision of two public goods: air quality and school quality. Income is assumed to be constant (\$50,000) in order to simplify the example. Panels *a* and *b* of the figure show how the three communities are positioned in geographic space and attribute space. Panel *c* uses the indirect utility function from equation (3.1) and the partitioning logic described in chapter 2 to define the region of preference space that is consistent with each location choice²⁷. For example, a household

²⁷ This example assumes $\beta = 2$, $\eta = -0.963$, $\nu = 0.75$, and $\rho = -0.01$.

with preferences anywhere in region A of the partition would maximize its utility by choosing to live in community A. Finally, panel *d* translates the partition of preference space into the average value for the marginal willingness-to-pay (per household) for a small air quality improvement in each community. It does this by assuming that preferences are uniformly distributed within each region of the partition.

Figure 3.2 replicates figure 3.1 after adding one additional community (D) to the choice set. Comparing the partitions depicted in panel *c* of each figure shows that adding community D diminishes the region of preference space assigned to households in A, B, and C. Notice that this changes conclusions about the MWTP. Thus, the two figures illustrate Samuelson's (1948) observation that consumers' preferences are identified by the choices they make and by the choices they could have made, but did not.

While expanding the study region may help to identify preferences, it may also bias estimated welfare measures if the region becomes so large that the free mobility assumption fails. For example, suppose that community D is a new housing subdivision and the annualized cost of moving to D from community A, B, or C is \$350.²⁸ Figure 3.3 contrasts the partition of preference space and the average MWTP when moving costs are ignored (panels *a* & *b*) with their counterparts when moving costs are incorporated (panels *c* & *d*). In panel *c* only the households with especially strong preferences for air quality—those with preferences in region D—will move into community D. Households with preferences in regions DB, DA, and DC will remain in communities B, A, and C. These households would move into D if it was costless to do so, but they are unwilling to pay the (effective) \$350 moving cost. Ignoring moving costs implicitly assumes that these households would move into D. Comparing panels *b* and *d* illustrates that this results in an upward bias on the estimated MWTP in community D and a downward bias on the MWTP in the other three communities.

²⁸ Moving costs are annualized in order to be consistent with annual income in the indirect utility function. While the physical costs of moving are a one-time expense, it seems reasonable to expect that the information and particularly the psychic costs may accrue over a longer period of time.

(B) *Intensive Margin: Defining Housing Communities & Labor Markets*

Once the study region has been defined, it must be subdivided into a set of housing communities and labor markets. As in part (A), the assumptions that underlie the empirical model provide some general guidelines for the subdivision but in practice the analyst is required to make decisions that can influence the resulting estimates. In the housing market, the key issue is the spatial aggregation of public goods. Recall that public goods are assumed to be homogeneous within each community. Therefore, by the law of one price, p_j should also be constant within each community since it represents the cost of access to public goods. For some public goods, the homogeneity assumption may suggest natural boundaries. School quality may vary discretely from school district to school district, and most of the variation in air quality may occur between air basins. In this case, communities could be defined using the smaller of the two spatial aggregation concepts. However, the homogeneity assumption seems less plausible for public goods defined by access to an amenity at a particular point in space. Distance to the city center, a national park, or the nearest beach will all vary within and between communities. For these and other access-based amenities, smaller communities would provide more precise measures of average distance.

Of course, incorrectly aggregating communities can bias the resulting welfare measures. For example, suppose the 4-community version of the study region in figure 3.2 represents the “true” choice set. Figure 3.4 shows the consequences of incorrectly aggregating communities A and B. The price of housing and provision of air and school quality in the new community AB is calculated as a simple average of their levels in communities A and B. This aggregation affects measures of the MWTP in every community. In particular, notice that the estimated MWTP in community AB (figure 3.4.d) is lower than it was in A or B (figure 3.2.d). Since AB has lower air quality than A, the logic of revealed preferences would suggest that the average person in AB will have weaker preferences for air quality than the average person in A. While AB has higher air quality than B, it is also more expensive. The higher price of housing in AB makes it more difficult to explain why households with strong preferences for air quality would choose AB over D.

Subdividing the study region into a set of labor markets parallels the subdivision into housing communities. The definition for a labor market is guided by the maintained assumption that the average wage paid to workers in each occupation is constant within a labor market and varies between labor markets. Likewise, as with housing communities, improperly defining labor markets may bias the resulting welfare measures.

Once housing communities and labor markets have been defined, the final step in preparing the choice set is to attach commute times to each job-house combination. With a large study region, some of these combinations may require multiple-hour commutes that relatively few households would consider as feasible location choices. This presents the analyst with a tradeoff similar to the one at the extensive margin. Including extreme commutes in the choice set will provide more information to identify preferences at the risk of introducing bias into the resulting welfare measures.

In an application of the model, data limitations are also likely to play an important role in how the choice set is defined. The application described in chapter 4 provides examples of some of the issues that may arise and also discusses how features of the data may be used to resolve some of the tradeoffs described above. For the remainder of this chapter, it will be assumed that the choice set has been defined such that public goods are homogeneous within each community, wages are constant within each labor market, and that it is costless to move between any two locations.

IV. The Estimator

The richness in the specification for utility poses two challenges for the inversion process underlying the revealed preference logic. It must account for the presence of unobserved public goods and it must account for heterogeneity in some of the structural parameters. Berry, Levinsohn, and Pakes (1995), Epple and Sieg (1999), and Bajari and Benkard (2005) have all developed estimators that address these challenges. However, their estimators require restrictions on the shape of the utility function and assumptions for the distribution of heterogeneous preference parameters that are not satisfied by the specification for indirect utility in (3.1). Specifically, to use the estimator developed by Berry, Levinsohn, and Pakes

would require households to have idiosyncratic “tastes” for individual locations and those tastes would have to satisfy the *iid* Type I extreme value distribution assumption. Epple and Sieg’s estimator requires the joint distribution of preferences and income to be lognormal and it requires households to have vertically differentiated preferences for public goods. Like the current model, the estimator developed by Bajari and Benkard does not place *ex ante* assumptions on the distribution of heterogeneous preference parameters. However, their approach requires the utility function to be linear and additively separable. Since the specification for utility in (3.1) violates the restrictions required to implement the existing structural estimators, a new approach must be developed. This section describes the new estimation framework.

The logic of the estimation process is described in figure 3.5. It can be decomposed into two stages. The first stage estimates the price of housing in each community (p_1, \dots, p_J) , and the homogeneous housing demand parameters (β, η, ν) . These results are treated as known constants during the second stage of the estimation, which recovers a composite unobserved public good for each community (ξ_1, \dots, ξ_J) , the homogeneous CES parameter (ρ) , and a partition of preference space for the heterogeneous parameters $A(\alpha, \gamma, \theta)$. All of the second stage parameters are estimated simultaneously, following the iterative process depicted in figure 3.5. Parts (A) through (E) of this section describe the estimation process and explain how each parameter is identified.

(A) *First Stage Estimation: The Price of Housing (p_1, \dots, p_J)*

Recall that housing is treated as a homogeneous commodity that can be consumed in continuous quantities. Under this assumption, the price of housing reflects the cost of consuming the public goods provided by each community. Of course, in practice housing is not homogenous. Its structural characteristics (e.g. bedrooms, bathrooms, sqft.) vary within and between communities, and these differences will be reflected in observable sale prices. This can be addressed if we are prepared to assume that the structural characteristics of housing enter the direct utility function through a sub-function that is homogeneous of degree

one and separable from the effect of public goods and the numeraire. Under this restriction, Sieg et al. (2002) demonstrate that the equilibrium locus of housing expenditures defined by a hedonic price function will be separable in the structural characteristics of houses and the effect of public goods, as shown in (3.7).

$$(3.7) \quad e_{j,n} = \bar{h}(h_{j,n}) \cdot p_j(g_{1,j}, \dots, g_{N-1,j}, \xi_j) .$$

The left side of the expression represents expenditures on house n in community j . The first term on the right side is a “quantity” index of housing that depends on a vector of structural characteristics ($h_{j,n}$). By condensing all the information about the structural characteristics of a house into a single number, the index provides an empirical analog to the concept of a homogeneous unit of housing. The second term represents the price of a homogeneous unit of housing in community j , which depends on the public goods it provides, observed and unobserved. Taking logs of (3.7) produces a version of the housing price hedonic model shown in (3.8).

$$(3.8) \quad \ln(e_{j,n}) = \ln[\bar{h}(h_{j,n})] + \ln[p_j(g_{1,j}, \dots, g_{N-1,j}, \xi_j)] + \mu_{j,n} .$$

Given a parametric form for (3.8) and data on housing transaction prices and their structural characteristics, the price of housing in each community can be recovered as a community-specific fixed effect.

(B) *First Stage Estimation: Housing Demand Parameters* (β, η, ν)

Having recovered housing prices, the next step is to use them along with data on housing expenditures and household income to estimate the homogenous housing demand parameters (β, η, ν). An individual household’s demand for housing can be derived from the indirect utility function as: $\bar{h}_i = \beta p_z^\eta y_i^\nu$ (as in equation [3.3]). Taking logs, equation (3.9) provides an expression for the Nth quantile in the housing demand distribution for community j .

$$(3.9) \quad \ln H_j^N = \ln(\beta) + \eta \ln(p_j) + \nu \ln y_j^N$$

Multiplying both sides of (3.9) by the price of housing produces the expression for housing expenditures in (3.10), where expenditures are assumed to be measured with error. The intercept in the demand for housing can be estimated together with the price and income elasticities by regressing quantiles of the distribution of annualized housing expenditures, e_j^N , on the price of housing and quantiles of the income distribution, y_j^N . While a single quantile is sufficient to identify the demand parameters, adding data on additional quantiles can increase the efficiency of the estimator.

$$(3.10) \quad \ln(e_j^N) = \ln(\beta) + (\eta + 1) \ln(p_j) + \nu \ln(y_j^N) + \varepsilon_j.$$

Since housing prices were estimated as fixed effects in a hedonic regression of (3.8), they will be measured with error. The observable public goods can be used as instruments for price. In addition, non-wage income can be used as an instrument for income, which will be endogenous due to spatial variation in wages. Assuming the error terms in (3.10) are uncorrelated across different quantiles of the distribution of income and expenditures, the quantiles can be stacked and the regression can be run using 2SLS.²⁹

Throughout the second stage of the estimation the first stage estimates are treated as known constants³⁰. To reduce notation in the following discussion, let δ represent the first stage results plus all the data on attributes of locations: $\delta = [\beta, \eta, \nu ; p, g, w, t]$.

²⁹ Alternatively, if the error terms are expected to be correlated across quantiles, the estimation could be performed using GMM or using SUR with restrictions on the parameters across equations.

³⁰ One alternative strategy would use the endpoints of a confidence interval on each parameter in (3.10) to place bounds on the second stage parameters. A second alternative would be to use the assumed distributions for the first stage parameters to generate distributions for $(\rho, \bar{\xi})$ in the second stage. A third approach would be to estimate $\beta, \eta, \nu, \rho, \bar{\xi}$ simultaneously in a GMM framework, noting that any additional identification of β, η, ν would come from the independence restriction used to identify ρ in the second stage.

(C) *Second Stage Estimation: Composite Unobserved Public Good* $(\bar{\xi}_1, \dots, \bar{\xi}_J)$

Communities are usually differentiated by some public goods that matter to households but are unobservable to the econometrician. Their influence is revealed in the data by the presence of seemingly inferior communities which, compared with other plausible locations, have higher housing prices and lower provision of every observed public good. Without accounting for unobserved public goods, the estimator cannot explain why a household would choose an inferior location. Fortunately, the way that households react to the supply of unobserved public goods helps to identify composite provision of unobserved public goods in each community.

If unobserved public goods influence households' location choices, they should also influence the price of housing. Under the maintained assumption that households have nonnegative preferences for public goods, the price of housing will be strictly increasing in unobserved public goods as in (3.11.a).³¹

$$(3.11.a) \quad \frac{\partial p(g_1, \dots, g_{N-1}, \bar{\xi})}{\partial \bar{\xi}} > 0 .$$

$$(3.11.b) \quad \bar{\xi} \perp g_1, \dots, g_{N-1} .$$

If (3.11.a) holds, the price of housing in each community that was recovered as a fixed effect in (3.8) should contain information about the provision of public goods in that community. More specifically, after controlling for the variation in the price index due to observed public goods, the remaining variation can be attributed to unobserved public goods. However, theory does not suggest a functional form for the relationship between p and $g_1, \dots, g_{N-1}, \bar{\xi}$. Importantly, the function need not be separable in observed and unobserved public goods³².

³¹ Recall from chapter 2 that Bajari and Benkard (2005) prove a hedonic price function for housing exists and is strictly increasing in $\bar{\xi}$ if utility satisfies differentiability, continuity, and nonsatiation in $\bar{\xi}$ and the numeraire (see equation [2.21]). These conditions are satisfied for the indirect utility function in (3.1).

³² The parameterization of the indirect utility function implies that MWTP for a change in an observed public good is a nonseparable function of the unobserved public good. Access to tennis courts, hiking trails, and other

Given this indeterminacy, the strategy used here is to impose the additional independence restriction in (3.11.b) which allows $\bar{\xi}$ to be recovered nonparametrically whether the price index is separable or nonseparable in the public goods.

When (3.11.a,b) hold, Matzkin (2003) implies that the quantiles of the distribution of the unobserved public good will equal the quantiles of the price distribution, conditional on observed public goods. This result is shown as (3.12).

$$(3.12) \quad F_{\xi}(\bar{\xi}_j) = F_{p|g=g_j}(p_j) = F_{p|g=g_j}[f(g_j, \bar{\xi}_j)].$$

A variety of methods can be used to map the price of housing in each community into its corresponding quantile in the distribution of prices, conditional on observed public goods. Examples include the kernel estimator in Matzkin (2003) and the local linear method used by Bajari and Benkard (2005). In either case, the estimated quantiles are a monotonic transformation of the unobserved characteristic itself since, assuming $\bar{\xi}$ has a continuous distribution, it can be normalized such that its marginal distribution is U[0,1]. This normalization implies $\bar{\xi} = F_{\xi}(\bar{\xi}_j)$.

The estimated values of $\bar{\xi}$ must permit the indirect utility function to explain every observed location choice, including the choice of seemingly inferior communities. This requires a certain degree of smoothness in the relationship between the price of housing and the unobserved public good³³. In practice, the minimum bandwidth that delivers this smoothness may exceed the bandwidth that would otherwise be chosen to address the bias/efficiency tradeoff from estimating (3.12). In the estimation, this is treated as a constraint on the bandwidth. The estimator starts with the “optimal” bandwidth. Then, if necessary, the bandwidth is increased until the estimator finds values for the heterogeneous

outdoor amenities (which are unobserved) may have higher value in locations with better air quality (which is observed). This relationship may or may not lead to nonseparability in the price index.

³³ This is a common feature of pure characteristics-based models such as Feenstra and Levinsohn (1995), Epplé, Peress, and Sieg (2005), and Bajari and Benkard (2005). Similarly, in mixed logit applications such as Berry, Levinson, and Pakes (1995) and Bayer, McMillan, and Reuben (2005) the idiosyncratic logit error terms “pick up the slack” in explaining choices.

parameters that justify every observed location choice. This process is shown in figure 3.5 by the arrows that link estimation of $\bar{\xi}$ to the process of partitioning preference space³⁴. The proof of theorem 1 in appendix A demonstrates the existence of some threshold bandwidth above which every location can be justified by nonnegative values of the heterogeneous parameters.

Figure 3.6 uses an example to illustrate how $\bar{\xi}$ can be estimated using a nonparametric kernel estimator. The sample data provided below the figure describe the price of housing and the levels of air quality and school quality in each of 20 communities. Without accounting for the presence of unobserved public goods, 9 of the communities appear to be inferior. For example, community 6 has a higher price of housing and lower levels of both public goods than community 5. Likewise, communities 9, 11, and 12 seem inferior relative to community 8. This problem can be addressed by estimating $F_{p|g=g_j}(p_j)$ nonparametrically from the data on housing prices, air quality, and school quality. Following equation (3.12), the resulting values can be interpreted as normalized measures for the level of unobserved public goods in each community. The dashed line in figure 3.6 depicts the values for $\bar{\xi}$ estimated using a multiplicative Gaussian kernel with a vector of bandwidths (q) chosen under the assumption that the underlying distributions of housing prices, air quality and school quality are each normally distributed. The corresponding values for $\bar{\xi}$ are reported in the q column of the table. Notice that communities 9, 11, and 12 are each assigned values for $\bar{\xi}$ that exceed the value assigned to community 8. Thus, they no longer appear inferior. However, community 6 is still inferior relative to 5 and community 19 is still inferior relative to 17. Therefore, some additional smoothing is required. The solid line in the figure depicts the values for $\bar{\xi}$ recovered by repeating the kernel estimation after increasing the bandwidths fivefold ($5q$). The corresponding values in the table confirm that no community is inferior.

³⁴ In contrast, in Epple and Sieg (1999) the assumed form of the joint distribution of income and preferences allows them to identify $\bar{\xi}$ from observed population shares.

(D) *Second Stage Estimation: CES Parameter (ρ)*

Given the data, the first stage estimates, and values for the unobserved public goods, households' location choices can be expressed as a function of the elasticity of substitution parameter, preferences for public goods, the opportunity cost of time, and unobserved job skill $(\rho, \alpha, \gamma, \theta)$. Unfortunately, observed location choices are not sufficient to simultaneously identify ρ and (α, γ, θ) without some knowledge of the relationship between preferences and income. Previous locational equilibrium applications have supplied this information by specifying a parametric form for their joint distribution (Epple and Sieg [1999]) or assuming they are independent for a subset of households (Epple, Peress, and Sieg [2005]). The later approach is used here.

All else constant, the interaction between ρ and non-wage income (\hat{y}) in the indirect utility function dictates how income shocks affect the desired combination of housing and public goods. This relationship can be inverted to identify ρ from the location choices made by households that are identical except for their non-wage income. Put differently, the observed stratification by income of (otherwise) identical households reveals the extent to which they substitute public goods with the private good component of housing.

Let $F_s(\alpha, \gamma, \theta)$ denote the distribution of the heterogeneous parameters for a subset of households, s , for which $F_s \perp \hat{y}$. Suppose this subset can be further divided into two groups with non-wage income \hat{y}_1 and \hat{y}_2 . The Gibbs algorithm outlined in part (E) can be used to partition preference space for each group. Sampling over the resulting partitions will produce two approximations to F_s , $\tilde{F}_{s,1}$ and $\tilde{F}_{s,2}$. These conditional distributions will equal the unconditional distribution (F_s) only when the partitioning process is performed at the true value of the CES parameter, $\rho = \rho_0$, as depicted in equation (3.13).

$$(3.13) \quad F_s(\alpha, \gamma, \theta) = \tilde{F}_{s,1}(\alpha, \gamma, \theta | \hat{y}_1, \rho, \delta, \bar{\xi}) = \tilde{F}_{s,2}(\alpha, \gamma, \theta | \hat{y}_2, \rho, \delta, \bar{\xi}), \text{ for } \hat{y}_1 \neq \hat{y}_2.$$

For other values of the CES parameter, $\rho \neq \rho_0$, the expression in (3.13) cannot hold. This

follows from the observation that the boundary indifference loci defining the partition of preference space are nonseparable in (ρ, \hat{y}) .³⁵ A movement in ρ away from its true value will distort the boundaries of the partition to a different extent for \hat{y}_1 and \hat{y}_2 , leading to predictions for $F_{s,1}$ and $F_{s,2}$ that differ from the true distribution and from each other: $\tilde{F}_{s,1} \neq \tilde{F}_{s,2} \neq F_s$. The estimator applies this logic to recover the value of ρ that minimizes the predicted difference between $F_{s,1}$ and $F_{s,2}$, as shown in equation (3.14).

$$(3.14) \quad \min_{\rho} \left\| \tilde{F}_{s,1}(\alpha, \gamma, \theta | \hat{y}_1, \rho, \delta, \bar{\xi}) - \tilde{F}_{s,2}(\alpha, \gamma, \theta | \hat{y}_2, \rho, \delta, \bar{\xi}) \right\|.$$

If location choices can be observed for s -type households at more than two income levels, the efficiency of the estimation may be improved by minimizing the difference between the predicted distributions for all pairwise combinations of income. In general, evaluating the objective function requires partitioning preference space at each of the $d = 1, \dots, D$ income levels and then sampling from those partitions to obtain $\tilde{F}_{s,1}, \dots, \tilde{F}_{s,D}$. This process must be repeated, updating ρ on each step, until the relevant convergence criteria are satisfied. Figure 3.5 depicts this iteration with arrows linking the estimation of ρ to the partitioning process.

(E) *Second Stage Estimation: Partitioning Preference Space* $A(\alpha, \gamma, \theta)$

Given the demand for housing, a value for the elasticity of substitution between public and private goods, and all the attributes that differentiate communities and jobs, location choices can be expressed as a function of preferences for public goods, the opportunity cost of time, and unobserved job skill. The partitioning process inverts this relationship, using the logic of revealed preferences to recover values for the heterogeneous parameters that rationalize observed location choices. This step of the estimation manifests Tiebout's logic that location choices reveal preferences.

³⁵ An expression for the boundary indifference loci is provided in equation (3.16).

The borders that delineate the partition are implicitly defined by the system of equations that arise from applying the boundary indifference condition (see equation [2.43]) to the indirect utility function in (3.1). This system is highly nonlinear. Consequently, the borders cannot be expressed analytically and when preference space exceeds two dimensions it is infeasible to solve for them numerically. Instead, the estimator recovers an approximation to the partition of preference space by sampling over it uniformly. Similar strategies have been used in the past by Feenstra and Levinsohn (1995) and Bajari and Benkard (2005).

As in Bajari and Benkard (2005), the sampling is done by a Gibbs algorithm that takes a large number of uniform draws from each region of the partition. Relative to their implementation the current problem poses three new computational challenges, all of which stem from the nonlinearity of the indirect utility function in (3.1). Before discussing these challenges, it is useful to briefly sketch the mechanics of the Gibbs sampler using a simple numerical example³⁶.

Figure 3.7 shows the partition of preference space from the 4-community example that was discussed in section III. Suppose we want to sample uniformly over region A of the partition but lack an analytical expression for its bounds. To start the Gibbs sampler, one must first locate a point somewhere in A. In the figure, the starting value is denoted by $*_0$. The first step is to condition on all but one coordinate and solve for bounds on the remaining coordinate. In the figure, this is done by conditioning on γ_{air} and solving for the bounds on α , which are 0.9 and 2.6. Use these bounds to take a random uniform draw. Suppose the result is $\alpha = 2.35$. From here, condition on $\alpha = 2.35$, solve for the bounds in the γ_{air} dimension, and take a random uniform draw on γ_{air} . In the figure, the new bounds are 0.0 and 0.4, and the new uniform draw is 0.13. Together, the two conditional uniform draws (2.35, 0.13) define the first unconditional draw from the region, $*_1$. This process can be repeated, using $*_1$ to find $*_2$ and so on. The result is a randomly chosen uniform distribution of points within region A of the partition.

There are three reasons why the nonlinearity of the indirect utility function

³⁶ Geweke (1996) provides a formal description of Gibbs sampling.

complicates the sampling process. First, the borders of each region in the partition are implicitly defined by the following system of nonlinear equations:

$$(3.15) \quad V_{i,z}(\alpha_i, \gamma_i, \theta_i | \hat{y}, \rho, \delta, \bar{\xi}) = V_{i,q}(\alpha_i, \gamma_i, \theta_i | \hat{y}, \rho, \delta, \bar{\xi}) \quad \forall q = 1, \dots, Z, \text{ where } q \neq z.$$

To solve for a single point on the boundary of region z requires solving for the roots of these $Z-1$ equations and then determining which of the roots defines an upper or a lower bound. This can be time-consuming. The second difficulty is that because the marginal utility of income depends on preferences, the partition of preference space depends on income. This means the entire sampling process must be repeated for every level of non-wage income for workers of each occupation. Thus, the number of nonlinear rootfinding problems that must be solved may be quite large. Third, the preference set that rationalizes any given location may be disconnected. Section V describes how the standard Gibbs sampling approach was adapted to address these computational challenges.

V. Computational Approach to Partitioning Preference Space

This section provides a more technical discussion of the Gibbs sampling algorithm. Part (A) starts with a step-by-step description of how the algorithm can be used to draw a uniform sample of points from one region of the partition of preference space, given a starting value in that region. This extends the version of the Gibbs sampler developed by Bajari and Benkard (2005) to the case of a non-linear utility function. However, the algorithm can not be guaranteed to converge unless the region of preference space is connected. Connectivity is only guaranteed in the special case where public and private goods are perfect substitutes (i.e. $\rho = 1$). Meanwhile, applications of Epple-Sieg type sorting models have unanimously rejected the perfect substitutability hypothesis. Therefore, part (B) of this section modifies the Gibbs sampler from part (A) to allow it to “jump” between disconnected sets that comprise a single region of preference space. This ensures the algorithm can converge when $\rho \neq 1$. Finally, part (C) describes how to find starting values to initiate the sampling

algorithm.

(A) *Applying the Standard Gibbs Sampling Algorithm*

Given a starting value in a particular region of the partition, the Gibbs algorithm will sample uniformly within that region. Let the starting value be defined by a point:

$(\alpha^0, \gamma_1^0, \dots, \gamma_{N-1}^0, \gamma_\xi^0, \theta_1^0, \theta_2^0)$. The algorithm iterates over the following steps:

- (1) Given $(\gamma_1^0, \dots, \gamma_{N-1}^0, \gamma_\xi^0, \theta_1^0, \theta_2^0)$, solve for the minimum and maximum bounds on α , and use those bounds to take a new draw, α^1 .
- (2) Given $(\alpha^1, \gamma_2^0, \dots, \gamma_{N-1}^0, \gamma_\xi^0, \theta_1^0, \theta_2^0)$, solve for the minimum and maximum bounds on γ_1 , and use those bounds to take a new draw, γ_1^1 .
- ⋮
- ⋮
- ⋮
- (N+3) Given $(\alpha^1, \gamma_1^1, \dots, \gamma_{N-1}^1, \gamma_\xi^1, \theta_1^1)$, solve for the minimum and maximum bounds on θ_2 , and use those bounds to take a new draw, θ_2^1 , resulting in $(\alpha^1, \gamma_1^1, \dots, \gamma_{N-1}^1, \gamma_\xi^1, \theta_1^1, \theta_2^1)$, the first new point in the region. Using this new point, return to step (1) and repeat.

If the algorithm is sampling over region z , the upper and lower bounds in each step belong to the set of boundary indifference points between z and every other location. The roots of equation (3.16) define the boundary points between location z and any alternative q .

$$(3.16) \quad {}_z\alpha_q = \frac{X_z(\theta) - X_q(\theta)}{[\bar{g}_q(\gamma)]^\rho - [\bar{g}_z(\gamma)]^\rho}, \quad \text{where}$$

$$X_z = \left[\exp\left(\frac{y_z^{1-\nu} - 1}{1-\nu}\right) \exp\left(-\frac{\beta P_z^{\eta+1} - 1}{\eta+1}\right) \right]^\rho,$$

$$\bar{g}_z = \gamma_1 g_{1,z} + \dots + \gamma_{N-1} g_{N-1,z} + \gamma_\xi \bar{g}_z, \quad \text{and} \quad y_z = \hat{y} + \theta_1 w_z (1 - \theta_2 t_z).$$

In each step of the Gibbs algorithm, the relevant upper (lower) bound on region z is the supremum (infimum) of the set of upper (lower) bounds defined by the roots of (3.16) for z and all $q \neq z$. In step (1) these $Z-1$ roots can be obtained by solving the analytical expression for $(\alpha | \gamma, \theta)$ defined in (3.16), which is not computationally demanding. However, in steps (2) through $(N+3)$, the roots of $(\gamma | \alpha, \theta)$ and $(\theta | \alpha, \gamma)$ are implicitly defined by (3.16) and cannot be expressed analytically. This means solving $(Z-1)$ nonlinear rootfinding problems on each step—a substantial computational burden. Fortunately, the total number of rootfinding problems can be reduced from $Z-1$ to 2 by using a discrete-continuous formulation of the objective function together with a bisection algorithm³⁷. This is done by defining the objective function to be dichotomous while the algorithm searches for the *locations* that define the supremum and infimum, and then redefining it to be continuous while the algorithm searches for the roots of (3.16). For example, while searching among the candidates for the supremum in step (2), the objective function equals 1 for values of γ_1 at which more than one other location provide higher utility, and -1 at all other values. When the algorithm finds a point where only one other location provides higher utility (the location that defines the supremum), the objective function switches to a continuous measure to solve for the root of (3.16). Put simply, first the algorithm finds the *location* that defines the supremum, then it solves for the *supremum*.

In general, three conditions are jointly sufficient to guarantee that a Gibbs sampling algorithm will converge (Geweke [1996]). The first is that each preference set is lower semicontinuous. If this condition fails to hold, the algorithm can get “stuck” at boundary points like the one shown in panel (A) of figure 3.8. To eliminate this possibility in the current problem, each region in the partition is treated as an open set, guaranteeing lower semicontinuity. The second condition is that each preference set is locally bounded. Economic theory, common sense, or structural restrictions can be used to specify prior upper and lower bounds on each preference parameter, ruling out the type of unbounded region

³⁷ While quasi-newton methods are often faster than the bisection approach, I found they were unable to handle the discrete component of the rootfinding problem.

displayed in panel (B) of figure 3.8. The final requirement for convergence is that the sample space has a connected support. Panel (C) shows an example where this condition fails. The standard Gibbs sampling algorithm will be unable to “jump” from one subset to another, in which case the result will vary depending on whether the starting value was located in A_{z1} or A_{z2} . The preference regions defined by (3.16) may be disconnected when $\rho \neq 1$. Fortunately, the structure of preferences also provides a way to induce the algorithm to jump between disconnected sets.

(B) *Modifying the Gibbs Sampler to “Jump” Between Disconnected Sets*

Set jumping is facilitated the homogeneity of the public goods index. To see this, first notice that the A_z regions are necessarily connected along a ray through the origin in weights space³⁸. A change in the magnitude of the weights can always be exactly offset by a corresponding change in preferences for public goods relative to private goods. Mathematically, the public goods index is homogeneous of degree ρ so that multiplying the γ 's by λ and dividing α by λ^ρ leaves utility unchanged, as in (3.17). Thus, if one point along a ray is contained in A_z , so is every other point on that ray.

$$(3.17) \quad V_z = \left\{ \alpha (\gamma_1 g_{1,z} + \dots + \gamma_\xi \bar{g}_z)^\rho + X_z \right\}^{\frac{1}{\rho}} = \left\{ \frac{\alpha}{\lambda^\rho} (\lambda \gamma_1 g_{1,z} + \dots + \lambda \gamma_\xi \bar{g}_z)^\rho + X_z \right\}^{\frac{1}{\rho}}.$$

Rays in weights space can be used to characterize set disconnections. To see this, consider the case where income is exogenous and three communities (i, j, k) are ranked by price such that $(p_i < p_j < p_k)$. For a given set of γ 's, there exists a region of preference space consistent with the choice of community j if and only if ${}_j \alpha_k > {}_i \alpha_j$. This condition can be rewritten as (3.18), using the results from (3.16).

³⁸ In weights space, relative preferences are constant along a ray through the origin. Movement along the ray simply represents a rescaling of the weights.

$$(3.18) \quad \frac{(\gamma_1 g_{1,j} + \dots + \gamma_\xi \bar{\xi}_j)^p - (\gamma_1 g_{1,i} + \dots + \gamma_\xi \bar{\xi}_i)^p}{(\gamma_1 g_{1,k} + \dots + \gamma_\xi \bar{\xi}_k)^p - (\gamma_1 g_{1,j} + \dots + \gamma_\xi \bar{\xi}_j)^p} > \frac{X_i - X_j}{X_j - X_k} .$$

The right hand side is a constant that depends on housing prices, income, and structural parameters. Because the left hand side may be a non-monotonic function of γ_n , equation (3.18) may fail along a ray through the origin that lies between two other rays that intersect A_j , disconnecting the preference set³⁹. Figure 3.9 illustrates this for a numerical example with two public goods. No value of α can ever rationalize the choice of community j for a household with relative preferences defined by ray r .

The homogeneity of the public goods index allows the Gibbs sampler to jump between disconnected sets by randomly rescaling the weights during the sampling process. For example, suppose the algorithm starts at the point ($*_0$) in figure 3.9 and finds the new coordinate $\gamma_{school} = 0.5$. At this new coordinate, the discrete-continuous bisection procedure may find the bounds $\gamma_{air} \in [0, 0.1]$ and choose the new point ($*_1$); it may find the bounds $\gamma_{air} \in [0.37, 0.5]$ and choose the new point ($*_2$); or it may find the bounds $\gamma_{air} \in [0, 0.5]$ and choose the new point ($*_3$). In the third case, new points are repeatedly chosen within $[0, 0.5]$ until a point within region z is located. In summary, not scaling the weights to sum to 1 during the sampling process ensures that the algorithm can move vertically and horizontally through A_j , allowing it to jump over disconnecting rays. After the sampling process ends, the resulting α 's and γ 's can be rescaled so that the γ 's always sum to 1.

(C) Finding Starting Values in each Preference Set

To start the Gibbs sampler, one must first locate a starting value in each region of the partition of preference space. Fortunately the structure of preferences can be manipulated to

³⁹ Notice that the left hand side of (3.18) is invariant to a rescaling of all the weights. This is another way to see that there cannot be a disconnection along a ray through the origin in weights space.

reduce the total number of starting values that must be obtained to J , one for each community. To see this, first notice that for retired households a γ – vector that rationalizes the choice of community j at income level y will rationalize that same choice at any income level. This is because an increase in exogenous income can always be exactly offset by a decrease in α . In other words, rescaling income leaves the right hand side of (3.18) unchanged. Furthermore, if a point (α^*, γ^*) rationalizes the choice of community j for retired households, the point $(\alpha, \gamma^*, \theta_1 = \varepsilon, \theta_2 = \varepsilon)$ will rationalize the same community for working households in each of the $k = 1, \dots, K$ job locations, where ε represents a sufficiently small, but positive, value.

Equation (3.19) defines an objective function that can be used to search over preference space for a starting value.

$$(3.19) \quad f = \max \{ V_{k \neq j}(\alpha, \gamma | y, \rho, \delta, \bar{\xi}) \} - V_j(\alpha, \gamma | y, \rho, \delta, \bar{\xi}).$$

Any optimization algorithm can be instructed to search over (α, γ) to minimize (3.19) until it finds a point (α^*, γ^*) where $f < 0$. However, this search process is complicated by the fact that the objective function tends to have multiple local optima. In practice, the multiplicity of local optima can make it very difficult to find starting values using local search methods such as the Nelder-Mead algorithm and quasi-newton approaches. In such cases, global optimization methods, such as the genetic algorithm *gloop_genetic*, are likely to be more effective.

Like many genetic algorithms, *gloop_genetic* is a randomized adaptive search that utilizes a user-influenced adaptation of three laws of natural selection: *survival of the fittest*, *reproduction*, and *mutation*. The algorithm begins by evaluating an initial population of points, uniformly chosen from the specified search space. Then the more successful points (i.e. those with lower objective function values) survive to reproduce and mutate, creating the next generation, while the less successful points die out. Over time the population of points will migrate toward the global optimum. Appendix B contains a more detailed description of the algorithm.

VI. Conclusions

This chapter developed an estimation strategy to recover an individual household's preferences for public goods, given maintained assumptions on the indirect utility function, the choice set, and the process generating the data. The discussion started by generalizing Epple and Sieg's (1999) specification for utility to allow households to have horizontally differentiated preferences for public goods and to allow workers to differ in their job skills. In the resulting dual-market framework, each working household faces a limited set of wage-commute options. They may be forced to choose between lower-amenity communities with cheaper housing and better access to high-paying jobs and communities with higher amenities, poorer access, and more expensive housing. The location choices made by households facing this tradeoff reveal features of their preferences. That said, the analyst's role in defining these tradeoffs should not be understated. Using a simple example, it was shown that conclusions about the demand for public goods can depend on the way the analyst chooses to define a "housing community", a "labor market", and the geographic scope of the study region. Chapter 4 demonstrates how data on migration and commuting patterns can be used to help define these concepts in a way that is consistent with the economic logic underlying the sorting framework.

Generalizing the sorting framework to allow a joint job-house choice required developing a new approach to estimation. The new econometric strategy can be decomposed into two stages. First, the spatial distribution of housing prices, income, and public goods is used to recover parameters that characterize the demand for housing. Then, conditional on the resulting demand function, the estimator partitions preference space to identify a set of values for the heterogeneous (preference and job skill) parameters that rationalize each observed location choice. A key feature of the estimator is that it uses the logic of set identification to distinguish the identifying power of structural restrictions on the indirect utility function from the identifying power of maintained assumptions about the distribution of preferences. Put differently, the set identification approach eliminates the need for a

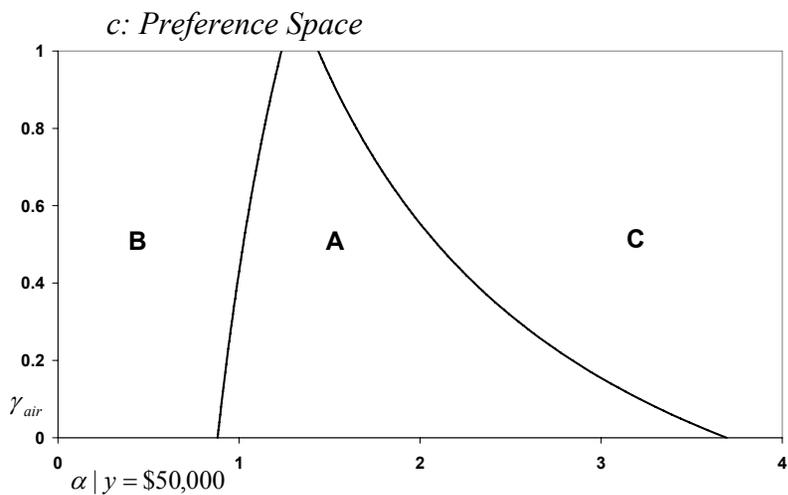
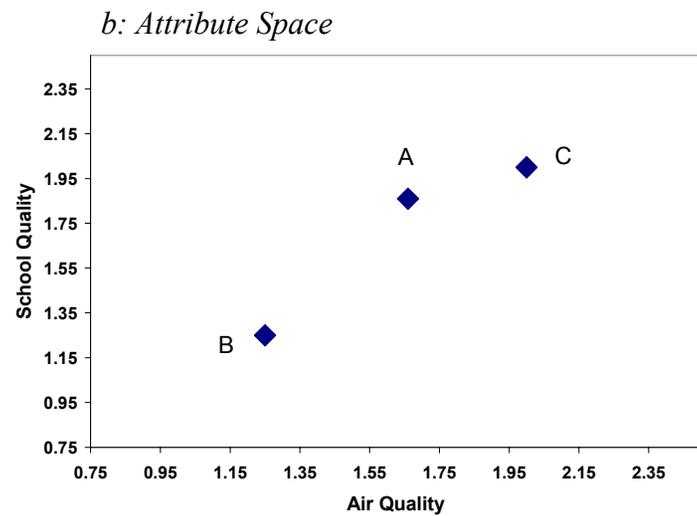
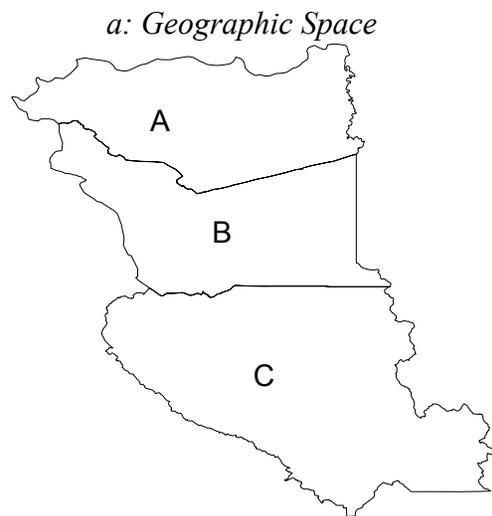
priori distributional assumptions about heterogeneous preference parameters.

Table 3.1 provides a comparison of the new estimator with the two closest alternatives—the sorting models developed by Epple and Sieg (1999) and Bayer, McMillan and Reuben (2005). The dual-market estimator has three advantages over the other two frameworks: (1) it accounts for labor market choices; (2) it eliminates the need for a priori distributional assumptions; and (3) it more fully utilizes the information in the hedonic price function to identify unobserved public goods. Notice that it nests Epple and Sieg’s framework as a special case when wages are fixed, preferences are vertically differentiated, income and preferences are joint lognormal, and the hedonic price function is not used to identify $\bar{\xi}$. Relative to Bayer et al., the disadvantage of the dual-market approach is that it fails to account for idiosyncratic tastes for individual locations. For example, the desire to live near family and friends may be an important determinant of where people choose to live. While the idea of idiosyncratic tastes is intuitively appealing, it is hard to imagine how they can be identified without distributional assumptions.

The next three chapters describe the data, results, and a policy application using the new dual-market framework to evaluate the importance of improved air quality in the San Francisco and Sacramento regions of Northern California.

Table 3.1: Comparison of Dual-Market Estimator to Alternative Sorting Models

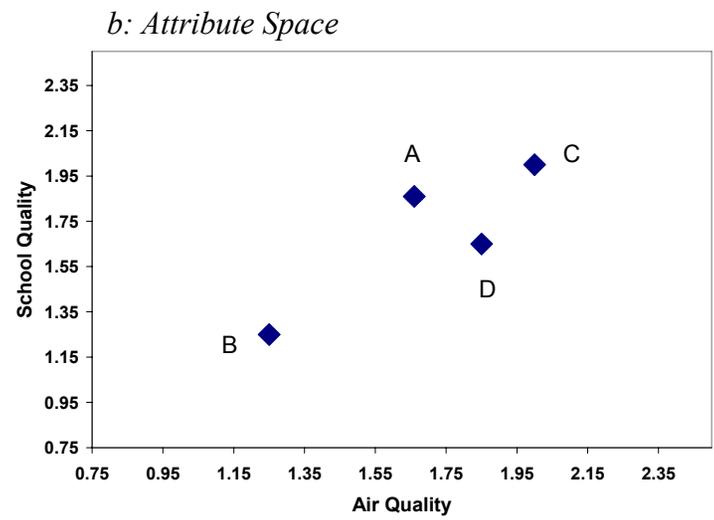
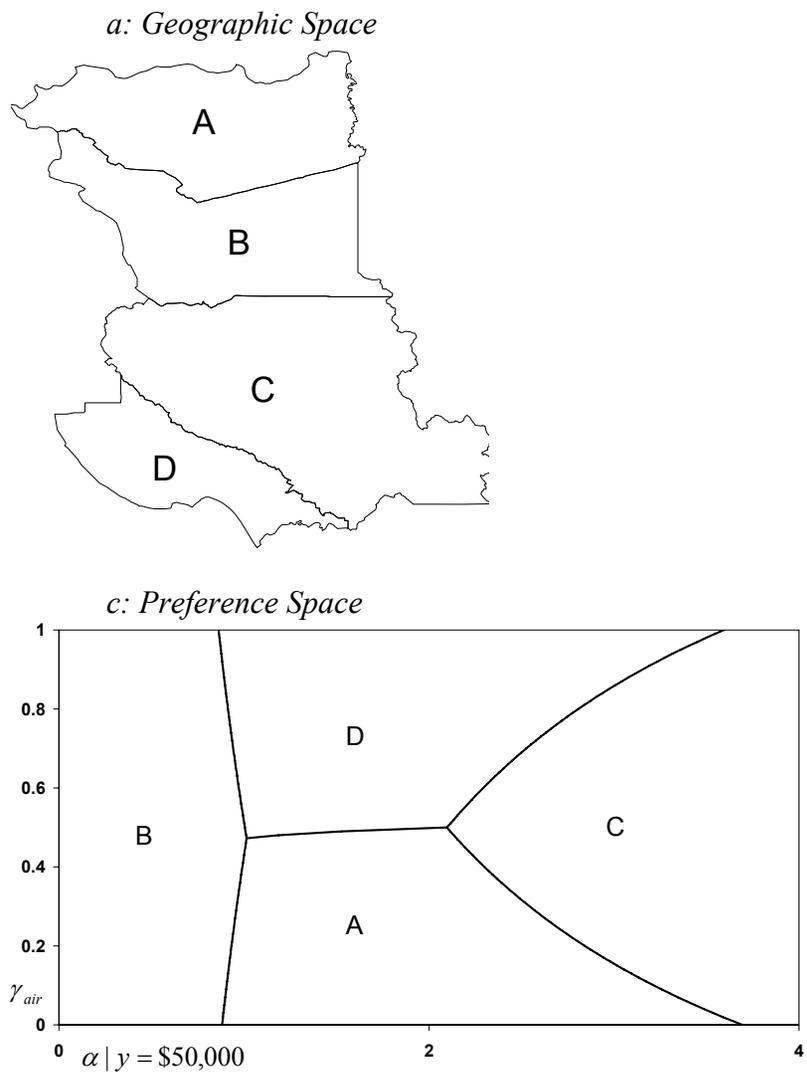
Category	Argument	Epple-Sieg (1999)	Dual-Market	Bayer et al. (2005)
Maintained Assumptions	Income	Fixed	Varies with job location	Fixed
	Free mobility	yes	Yes	yes
	Set of public goods	Discrete	Discrete	Discrete
	Set of housing characteristics	Continuous	Continuous	Discrete
Restrictions on Preference Heterogeneity	Preferences for observed characteristics	Vertical	Horizontal	Horizontal
	Preferences for unobserved characteristics	Vertical	Vertical	Horizontal
	Distributional assumptions	$F(\alpha, y) \sim \text{lognormal}$	None	$\varepsilon \sim \text{iid logit}$
Restrictions on Substitution Possibilities	Public goods w/ public goods	Perfect Substitutes	Perfect Substitutes	Perfect Substitutes
	Public goods w/ housing	CES utility \rightarrow anywhere from perfect substitutes to perfect complements	CES utility \rightarrow anywhere from perfect substitutes to perfect complements	Perfect Substitutes
	Substitutes per location (J total locations)	2	Between 2 and $J-1$	$J-1$
Source of Identification for Unobserved Public Goods ($\bar{\xi}_j$)		$F(\alpha, y) \sim \text{lognormal}$	Hedonic price function	$\varepsilon \sim \text{iid logit}$



d: MWTP for Improved Air Quality

Community	Air Quality	School Quality	Price of Housing	Average MWTP (per household)
A	1.66	1.86	1.26	131
B	1.25	1.25	1.00	97
C	2.00	2.00	1.50	382

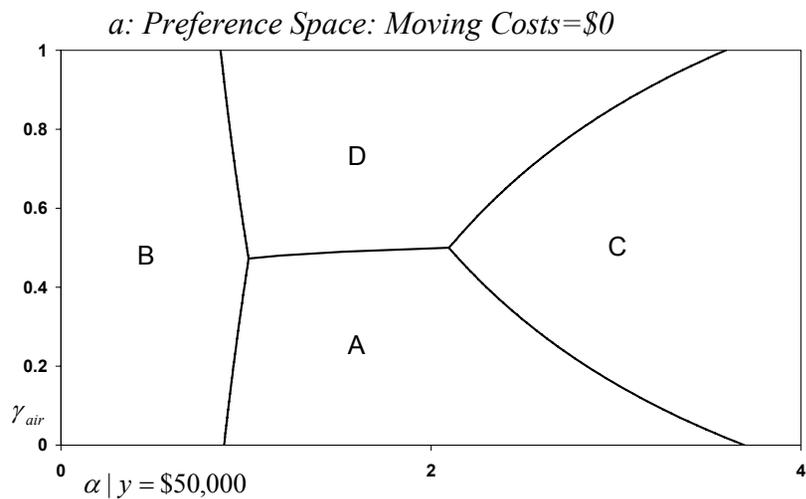
Figure 3.1: Three Communities in Geographic Space, Attribute Space, and Preference Space



d: MWTP for Improved Air Quality

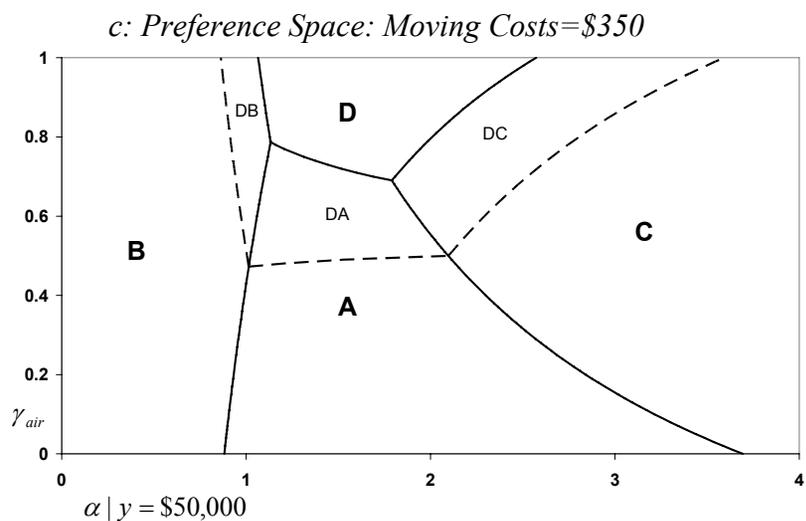
Community	Air Quality	School Quality	Price of Housing	Average MWTP (per household)
A	1.66	1.86	1.26	87
B	1.25	1.25	1.00	80
C	2.00	2.00	1.50	358
D	1.85	1.65	1.25	368

Figure 3.2: Four Communities in Geographic Space, Attribute Space, and Preference Space



b: MWTP for Improved Air Quality: Moving Costs=\$0

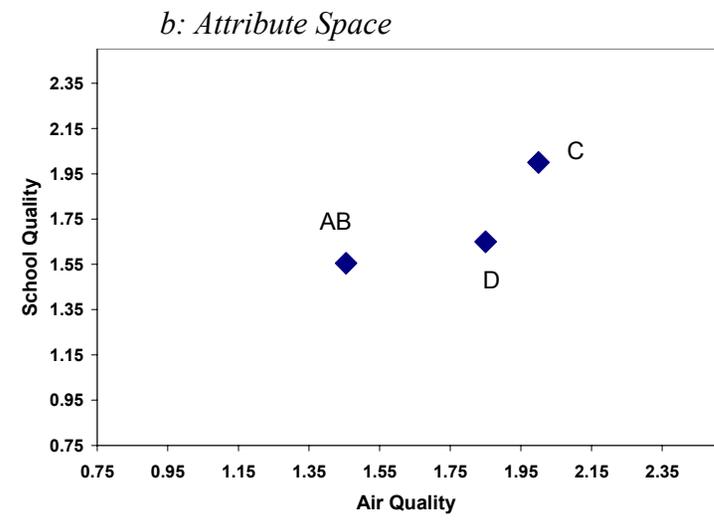
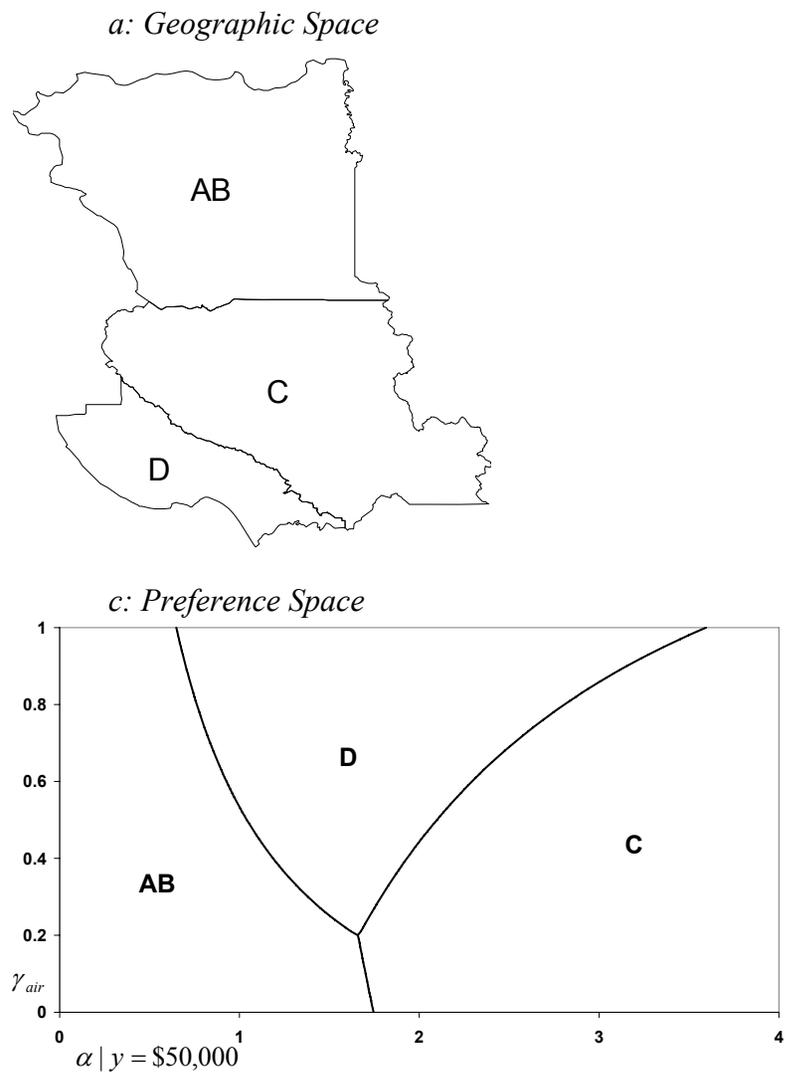
Community	Air Quality	School Quality	Price of Housing	Average MWTP (per household)
A	1.66	1.86	1.26	87
B	1.25	1.25	1.00	80
C	2.00	2.00	1.50	358
D	1.85	1.65	1.25	368



d: MWTP for Improved Air Quality: Moving Costs=\$350

Community	Air Quality	School Quality	Price of Housing	Average MWTP (per household)
A	1.66	1.86	1.26	111
B	1.25	1.25	1.00	94
C	2.00	2.00	1.50	383
D	1.85	1.65	1.25	354

Figure 3.3: Four Communities in Preference Space with and without a \$350 Cost of Moving to Community D



d: MWTP for Improved Air Quality

Community	Air Quality	School Quality	Price of Housing	Average MWTP (per household)
AB	1.46	1.56	1.13	59
C	2.00	2.00	1.50	284
D	1.85	1.65	1.25	332

Figure 3.4: Aggregating Communities A and B from Figure 3.2

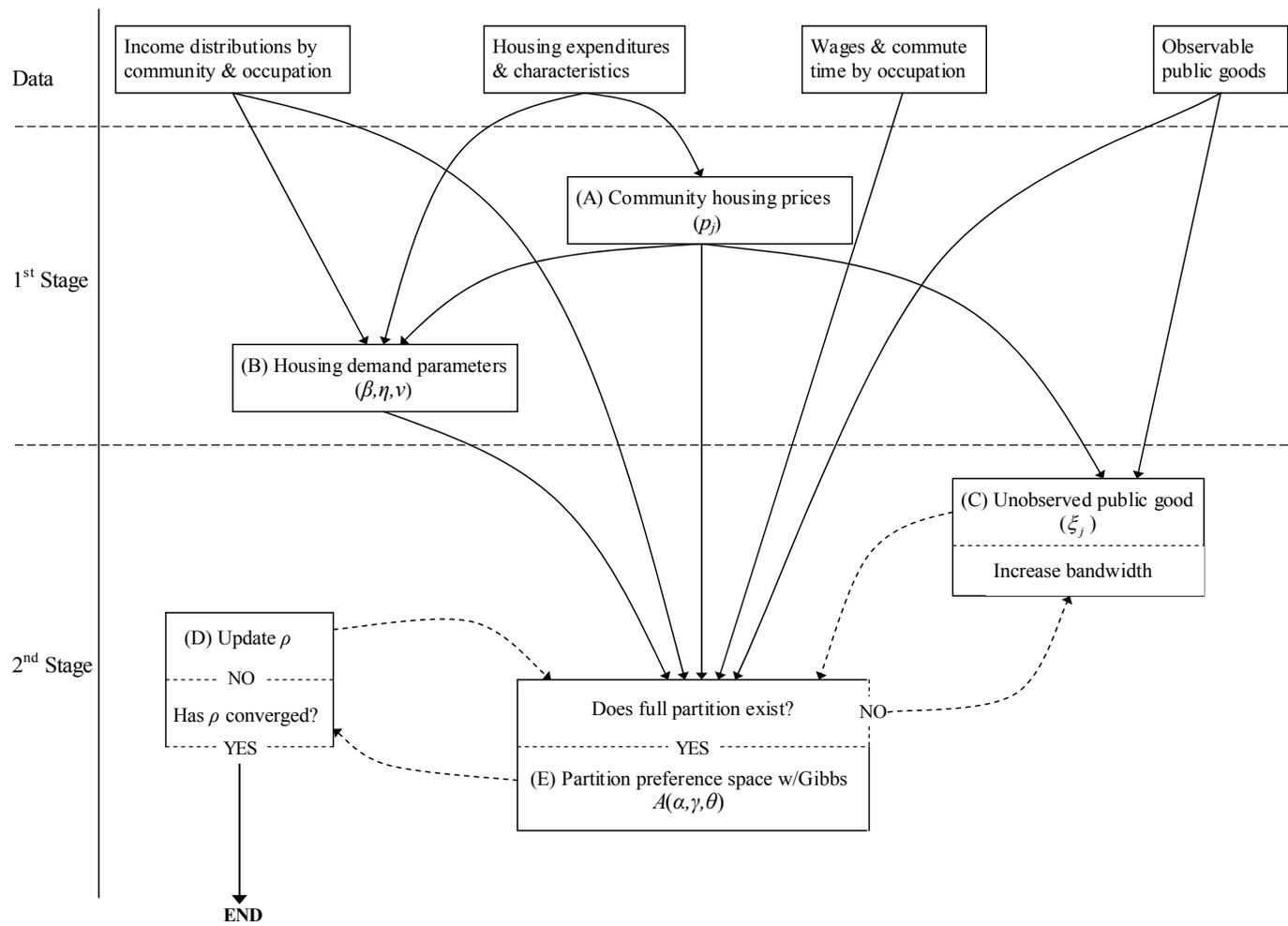
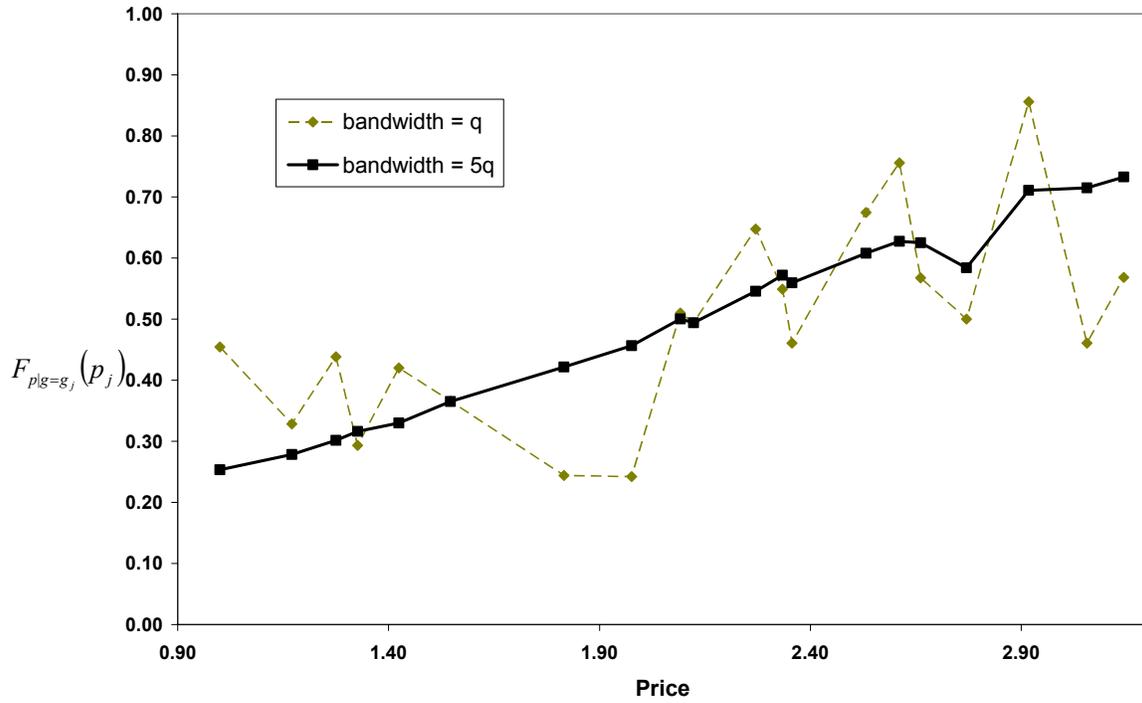


Figure 3.5: Overview of the Econometric Model



Community	Price	Air Quality	School Quality	ξ		Superior Community		
				q	5q	$\xi = 0$	q	5q
1	1.00	0.77	0.70	0.45	0.25	--	--	--
2	1.17	1.04	0.63	0.33	0.28	--	--	--
3	1.28	1.04	0.65	0.44	0.30	--	--	--
4	1.33	0.90	0.73	0.29	0.32	--	--	--
5	1.42	1.05	0.76	0.42	0.33	--	--	--
6	1.55	0.98	0.75	0.37	0.37	5	5	--
7	1.82	0.95	0.97	0.24	0.42	--	--	--
8	1.98	1.04	1.00	0.24	0.46	--	--	--
9	2.09	0.96	0.89	0.51	0.50	8	--	--
10	2.12	0.78	1.29	0.49	0.49	--	--	--
11	2.27	0.97	0.92	0.65	0.55	8	--	--
12	2.33	0.81	0.94	0.55	0.57	8	--	--
13	2.36	1.08	0.97	0.46	0.56	--	--	--
14	2.53	1.06	0.94	0.67	0.61	13	--	--
15	2.61	1.05	0.96	0.76	0.63	13	--	--
16	2.66	1.14	1.13	0.57	0.63	--	--	--
17	2.77	1.03	2.03	0.50	0.58	--	--	--
18	2.92	0.92	0.93	0.86	0.71	17	--	--
19	3.06	0.99	1.31	0.46	0.72	17	17	--
20	3.14	0.98	1.35	0.57	0.73	17	--	--

* This example used a multiplicative Gaussian kernel. q represents the bandwidth that would be optimal if the underlying distribution of each variable were normal

Figure 3.6: Using a Kernel Estimator to Recover the Unobserved Public Good (ξ) *

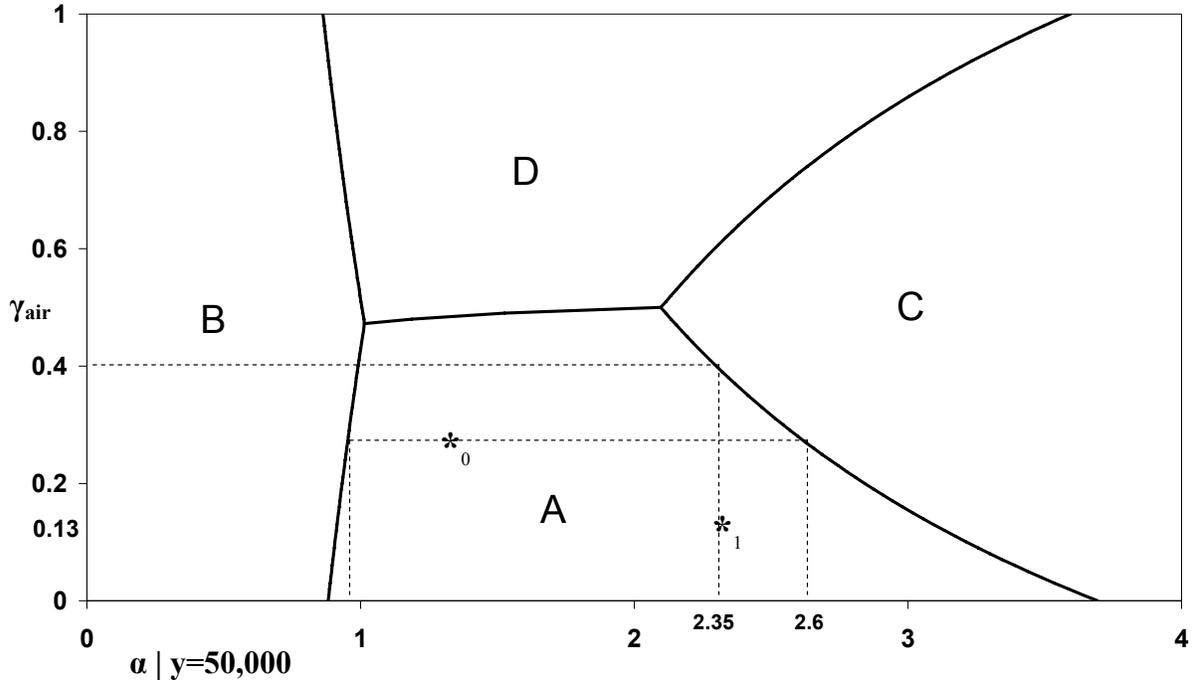


Figure 3.7: Gibbs Sampling Algorithm

(A) Not lower semicontinuous (B) Not locally bounded (C) Not connected

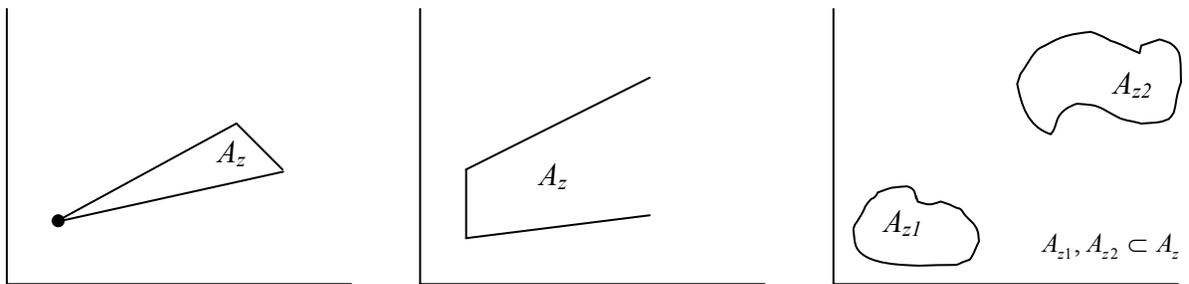


Figure 3.8: Convergence Problems with the Gibbs Sampler in Region A_z of the Partition

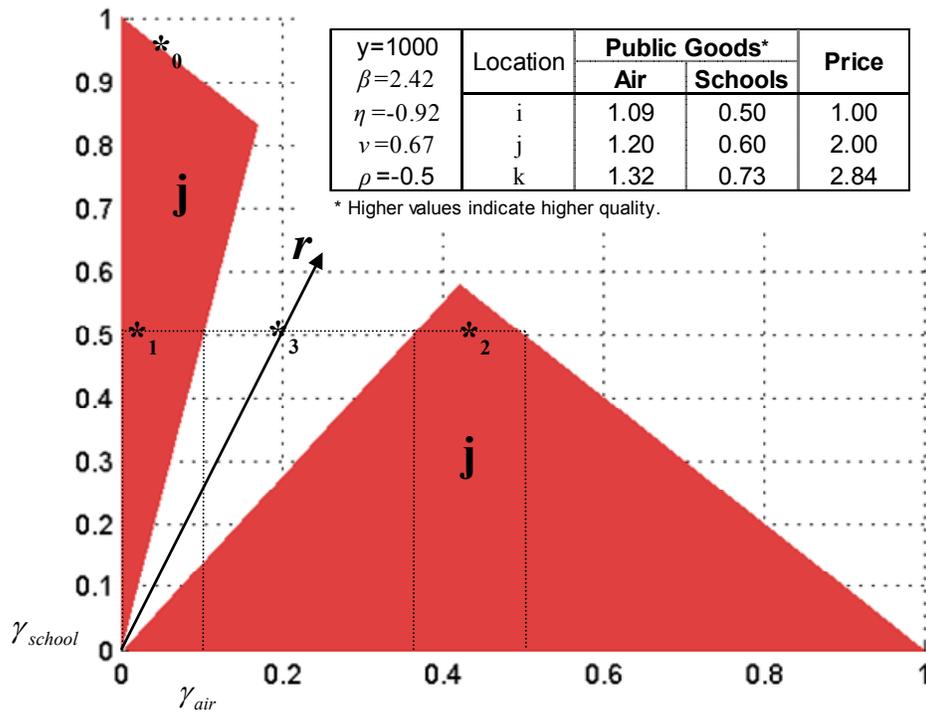


Figure 3.9: $A_j(\alpha, \gamma)$ is Disconnected

Chapter 4: Choice Set and Data

I. Introduction

This chapter describes the data that are used to estimate the empirical model for households in Northern California's two largest population centers: the San Francisco and Sacramento Consolidated Metropolitan Statistical Areas. While their major business districts are 100 miles apart, migration and commuting patterns support the decision to combine both into a single integrated region. This region provides a spatial landscape consistent with the development of the dual-market sorting model in chapter 2. That is, for a relatively small geographic area, the San Francisco-Sacramento region has tremendous diversity in housing prices, provision of public goods, and job opportunities. The diversity of opportunities in both the housing and labor markets creates the potential to learn more about households' preferences for public goods from observing a joint job-house choice than could be learned from housing choices alone.

Section II begins by summarizing conceptual issues that stem from recognizing the model as a data generating process. This provides a template for how the subsequent sections are organized. Section III describes the study region and how it was subdivided into a set of housing communities and work destinations. Each job-house combination in the region is differentiated by attributes of the housing community (housing prices, air quality, and school quality) and attributes of the work destination (wages and commute times). Section IV describes the data that provide values for these attributes in each location.

After defining the choice set, data must be collected on the job and house choices, incomes, and occupations of households throughout the study region. Unfortunately, these variables cannot be observed simultaneously in publicly available data sets. Section V describes the process that was used to generate an approximation to the income distribution for households in each location, by occupation. Finally, Section VI concludes with some discussion of how the resulting data set compares to the information used in related studies.

II. Generating the Data: An Overview

The empirical model relies on the logic that a household's location choices reveal its

preferences for the attributes of that choice. In this model a “location” is defined as the combination of a housing community and a labor market. The house component is associated with a set of local public goods that are purchased through the price of housing⁴⁰. The job component is associated with a wage rate which, in turn, feeds back into the demand for housing through its affect on income. Together, each job-house combination is associated with a unique commute time. Given values for these attributes, what is actually revealed by a household’s location choice will depend on the set of alternative locations that household could have chosen, but did not. It is the analyst’s job to define this choice set. How this is done will affect what is ultimately learned about preferences, as the discussion in chapter 3 demonstrated.

To estimate the empirical model, each household’s location choice must be observed along with the attributes that differentiate the locations comprising the choice set. There are three steps to developing the information necessary to implement this logic. First, the study region must be chosen and subdivided into housing communities and labor markets, and the set of all possible job-house locations must be defined. Second, data on the following attributes must be attached to each location: the price of housing, the levels of the public goods, the wage rates by occupation, and the commute time relevant to each area. Third, the number of households who choose each job-house location must be observed. This data-generating process integrates the type of micro data typically used in hedonic property value models with the type of aggregate data typically used in differentiated product models of demand in the empirical IO literature.

The previous chapter discussed how sorting theory provides some restrictions on the way the choice set is defined. First, households must be freely mobile within the study region. While provision of public goods and job opportunities can change over time, households’ location choices are observed at a single point in time. In order for this “snapshot” to accurately reveal preferences, households must have moved to the location with their preferred bundle of attributes. Therefore, the geographic boundaries of the study

⁴⁰ The “price of housing” refers to the price of a homogeneous unit of housing, as discussed in the previous chapters. Of course, the sale price of a home conveys a set of structural characteristics (e.g. bedrooms, bathrooms) along with location-specific public goods.

region need to be selected in a way that minimizes the extent to which the free mobility assumption is violated. Data on migration and commuting patterns can help to guide this process. Sorting theory also requires public goods to be homogeneous within each community and wage rates to be homogeneous within each labor market. The definitions for a “housing community” and a “labor market” must be selected in a way that balances this conceptual restriction with limitations on data availability.

After defining the choice set, values for the location-specific attributes must be assigned to each choice. For communities, this requires collecting data on housing prices and public goods. While all households are assumed to choose from the same set of communities, they face different job options. Each household is assumed to sort on the basis of the job options faced by its primary earner, which depend on that individual’s occupation. Thus, an “occupation” must be defined in such a way that each worker’s occupation can be observed and a wage rate for each can be attached to every work destination. Finally, commute times must be assigned to each location in the choice set.

The final step in preparing the data is to observe incomes, occupations, and locations for households throughout the study region. The challenge is to find a reasonable set of assumptions that make it possible to use publicly available data sets derived from the Census to approximate their joint distribution.

III. Defining the Choice Set

The San Francisco and Sacramento Consolidated Metropolitan Statistical Areas (CMSA), together, contain 14 counties and about 9 million people, roughly 25% of the state’s population and 3% of the U.S. population. While the two CMSAs are adjacent, their major business districts are about 100 miles apart—far enough to prohibit widespread commuting, but close enough that most households could move from one to the other without alienating friends and family, or having to readjust to a dramatically different environment. The proximity between these regions is also apparent in data on recent movers. Between 1995 and 2000, San Francisco was the top destination CMSA for people moving out of the

Sacramento area. Likewise, San Francisco was the top origin of people who moved into Sacramento.

While physical proximity and migration patterns suggest that San Francisco and Sacramento can be treated as part of the same choice set, it is less obvious how to define the external boundaries of the study region and how to subdivide it into a set of mutually exclusive locations. These decisions need to be evaluated because they will ultimately influence welfare measures. Part (A) of this section presents data on mobility that support the decision not to include other nearby metropolitan areas in the choice set. Then part (B) describes the logic that was used to subdivide the San Francisco-Sacramento region into a set of housing communities and work destinations.

(A) Using Evidence on Mobility to Define the External Boundaries of the Study Region

The conceptual discussion of the choice set in chapter 3 emphasized the connection between the definition of the study region and assumptions about household mobility. To recap, for observed location choices to reveal preferences, households must be able to move freely within the region. Furthermore, mobility outside the region must be restricted. This suggests the ideal study region would be relatively small and geographically isolated from alternative locations. The San Francisco-Sacramento area roughly fits this profile.

Land use, migration, and commuting patterns support the decision not to include other nearby metropolitan areas in the choice set. Figure 4.1 shows that, other than Redding, all of California's metropolitan areas form a contiguous region, with San Francisco and Sacramento bordering five other MSAs: Stockton, Modesto, Salinas, Merced, and Yuba. These areas are much less urbanized than San Francisco and Sacramento. They consist mainly of farmland and have fewer job opportunities, with a much larger share of the jobs in agriculture (table 4.1). The table also shows migration patterns between 1995 and 2000. During that period, there was little emigration from San Francisco and Sacramento into the adjacent MSAs (1.3% of the population) relative to the number of people who moved within the region (7.6%). Furthermore, households who live in San Francisco and Sacramento also

tend to work there. Only 1.5% of the workforce living in the two CMSAs reported commuting to a job outside the region during 1999 (Census county-to county worker flow files). Thus, Stockton, Modesto, Salinas, Merced, and Yuba seem less a part of Northern California's major population centers and more a part of the state's agricultural production regions⁴¹. One could even characterize the San Francisco-Sacramento region as an urban area surrounded by farmland, evoking the landscape metaphor Starrett (1981) used to motivate his internal capitalization logic.

Because the San Francisco-Sacramento area is relatively small, it seems reasonable to treat its households as being freely mobile. This assumption is reinforced by the observation that 7.6% of the population did move within the region between 1995 and 2000. The analytical model in the previous chapter decomposed moving costs into three components: *physical*, *information*, and *psychic*. Evidence from migration studies suggests the later two represent the majority of total moving costs. It seems reasonable to assume that information and psychic costs are minimal for moves within the San Francisco-Sacramento area. One can drive between the furthest points of the region in four hours, making it easy for recent movers to make "day-trips" back to visit family, friends, and a familiar environment, or for prospective movers to learn about the attributes of alternative locations. This leaves only the physical cost of moving⁴². Although physical costs are ignored in the present analysis, they are quantifiable in principle. Figure 3.3 illustrated how moving costs could be incorporated into the partitioning stage of the estimation if data on the number of recent movers were also available.

⁴¹ The counties that comprise the Stockton, Modesto, Merced, and Salinas MSAs are all among the top-ranked California counties in terms of farm-gate production value: Monterey (3), Merced (5), San Joaquin (6), Stanislaus (8) (California Agricultural Statistics Service). In comparison, the top-ranked county within the study region is Sonoma (16).

⁴² California's tax policy adds to these costs. Passed in 1978, Proposition 13 set the maximum property tax at 1% and limited annual increases in assessments to the minimum of a property's increase in market value, and 2%. However, when a property is sold it is reassessed at its full market value. To the extent that market values increase by more than 2% annually, households are effectively taxed for moving to a property of comparable value. While this constraint is potentially important, past evidence has suggested that its magnitude is small. O'Sullivan et al. (1995) report that with average annual inflation of 13%, Proposition 13 increases the time between moves by only 2 months for the average household.

(B) *Subdividing the Study Region into a Set of Job-House Locations*

Having defined the study region, the next step is to subdivide it into a discrete set of locations. This process requires a definition for “housing communities” and “labor markets”. The resulting set of job-house combinations contains some locations that would require unrealistic commutes. Data on commuting patterns are used to identify these locations and remove them from the choice set.

School districts were used to divide the study region into 122 housing communities. These are primarily unified school districts, although exceptions were made for primary and secondary districts that do not belong to a unified district and for the city of San Francisco which was divided into 11 supervisorial districts. Admittedly, a school district is an arbitrary delineation for measuring provision of public goods other than school quality. Yet they are regularly used in empirical applications because they tend to be consistent with the underlying sorting theory as well as limitations on data and computing power (e.g. Epple and Sieg [1999], Sieg et al. [2004], Smith et al [2004]).

Recall that sorting theory requires the provision of public goods to be homogenous within each individual housing community. It seems plausible that this restriction would be approximately satisfied for both air quality and school quality when communities are defined as school districts. While school quality may vary within a district, parents usually have the option to send their children to any public school within the district where they live. This makes provision of school quality effectively homogeneous within each district.

In the case of air quality, there is no obvious geographic concept that satisfies the homogeneity restriction. The California Air Resources Board divides the San Francisco-Sacramento region into two *air basins* that differ in their geographic and meteorological conditions. However, there is substantial variation in measures of air quality within each basin (e.g. ozone, particulate matter). Moreover, some of this variation is reported regularly by local news stations and regional newspapers. Composite measures of air quality such as EPA’s Air Quality Index are reported for individual counties on a daily basis. It is also possible that households perceive some within-county variation, although it is not typically reported by the media. On average, each county contains 10 school districts. They are

conveniently sized to measure variation in air quality at the highest level of spatial resolution available in the data. The study region contains 210 air quality monitoring stations, an average of 1.7 per school district. It seems unlikely that households would be able to perceive variation in air quality within a school district unless that variation was associated with a feature of the natural landscape such as the coast, a stationary source of emissions such as an oil refinery, or a quasi-stationary source such as a freeway.

In general, subdividing school districts to form smaller communities offers little potential to learn more about the demand for the observable public goods. If school districts were subdivided, communities within the same school district would be assigned identical values for school quality and virtually identical values for air quality by virtue of the spatial distribution of monitoring stations. Therefore, the additional variation in the price of housing stemming from the subdivision would be attributed to the presence of unobserved public goods during the estimation. This illustrates the general point that data limitations tend to constrain the definition for the choice set in applications of sorting models.

The San Francisco Unified School District is a special case. It represents the entire city of San Francisco—10% of all the households in the study region—and has a complex enrollment process. Parents can request that their children be enrolled in any public school in the city⁴³. Then the district assigns each student to a school based partly on proximity and partly on diversity goals. This assignment process introduces uncertainty into the link between location choice and public school quality. For example, in 2004 only 62% of students were assigned to the first school they requested. Despite the fact that expenditures per student are equalized across schools, average scores on standardized tests vary widely across the city, giving parents who live near low quality schools an incentive to request that their children be sent elsewhere. In addition to variable school quality, San Francisco contains neighborhoods that differ in their provision of local public goods including crime, culture, and microclimate. Given its size and diversity, the city of San Francisco was subdivided into 11 communities according to its supervisorial districts. While there is no guarantee that children living in a supervisorial district will attend its schools, proximity to

⁴³ Parents can provide the district with a ranked list of up to 7 schools.

the schools certainly increases that probability.

Subdividing San Francisco leaves Oakland Unified School District as the new largest community with 141,000 households. The smallest community is Point Arena Joint Union High with only 850 households and the median community has approximately 20,000 households⁴⁴. Figure 4.2A shows the 122 communities that comprise the housing component of the choice set.

Labor markets are defined as Primary Metropolitan Statistical Areas (PMSA). The Census Bureau describes a PMSA as “a large urbanized county or cluster of counties...that demonstrate very strong internal economic and social links, in addition to close ties to other portions of the larger [CMSA] area” (Census, 2000, Appendix A-16). This description seems consistent with the conceptual notion of a labor market as a region where the “law of one price” would require two workers with identical skills to face identical job options. Treating PMSAs as distinct labor markets is also common practice in applications of the interregional hedonic model (e.g. Roback [1982]; Blomquist et al. [1988]).

Figure 4.2B shows how the San Francisco-Sacramento area is subdivided into eight PMSAs and the density of Census tracts in Figure 4.2C illustrates that the population within these regions is mostly concentrated around the San Francisco Bay and the city of Sacramento. Rush-hour traffic in these population centers is heavy enough that commuting times can be substantial even between adjacent PMSAs. Table 4.2 shows that while there is moderate commuting between a few of the PMSAs, most workers live and work in the same one, and there is almost no commuting between the San Francisco and Sacramento CMSAs.

With 122 housing communities and 8 work destinations, the choice set contains 976 possible job-house combinations. However, many of these would require workers to drive over 100 miles to go to work. Not surprisingly, workers rarely choose to live so far from their job. For example, only 8 workers report commuting between the San Jose PMSA and Sacramento City Unified School District, a 120 mile drive in each direction. Just as migration patterns were used to define the external boundaries of the study region, commuting patterns were used to reduce the set of location choices within the region. The

⁴⁴ Point Arena Joint Union High has a small population because it is located in the rural part of Sonoma County. In terms of acreage, it is roughly twice the size of the Oakland Unified school district.

criterion used to define an admissible location is that it must account for at least 500 working households (0.02% of the working population). This rule reduces the choice set to 268 locations, effectively excluding multiple-hour commutes between opposite ends of the study region and most commuting between the two CMSAs.

The final choice set is defined by 268 job-house combinations that together account for 99% of working households in the San Francisco-Sacramento region. For retired households, the choice set simply contains the 122 housing communities. Each household is assumed to have chosen the location that maximizes its utility, given the job opportunities faced by its primary earner and its preferences for the public goods that differentiate communities. The next step in preparing the data is to attach values for these attributes to each location.

IV. Defining the Attributes that Differentiate Location Choices

Households are assumed to derive utility from the public goods in their home community and to pay for them through housing prices. Recall that the estimator treats the price of housing and the supply of public goods as being constant within each community. The first three parts of this section describe how data on (A) prices, (B) air quality, and (C) school quality were collected and transformed into community-specific measures. Then part (D) discusses other public goods that could have influenced households' location choices but are not directly observed.

Expanding the model to include the labor market means that data must be collected on the attributes that differentiate job locations. In the empirical model, job opportunities are differentiated by commuting time and by the wages paid to workers in each occupation. Part (E) explains how commute times were attached to each of the 268 job-house locations. Finally, part (F) describes the data used to define a set of wage rates in each of the 8 PMSAs.

(A) Individual Housing Transactions

The empirical model treats housing as a homogeneous good that can be consumed in

continuous quantities in each community at a “per/unit” price. Recall that chapter 3 described a procedure to translate data on individual housing sales into an index of these per/unit prices. The basic idea is to control for within-community variation in structural characteristics by using a hedonic regression to remove their effect on housing prices. The remaining community-specific fixed effects can be treated as a price index for the bundle of public goods. Estimating such an index requires a large number of observations on housing sales in each community.

Data on the sale price and structural characteristics of 2.9 million housing transactions in the San Francisco-Sacramento region were purchased from the commercial vender *DataQuick*, which originally collected the data from the Assessor’s office of each county. For each of the 14 counties, the data contain housing sale prices, dates, and addresses. All but one county also report a core set of structural characteristics: # bedrooms, # bathrooms, year built, square feet for the dwelling, and square feet for the lot. Yolo County is the exception. The data for Yolo include prices, dates, and addresses, but lack information on the other structural characteristics. Additional sources were used to supplement this information. Structural characteristics for 1082 houses sold in Yolo between 1998 and 2005 were acquired from the *Sacramento Bee* newspaper’s online real estate archives. Similar data were obtained from the *Metrolist* multiple listing service webpage for another 815 homes listed for sale in October, 2005. For each of these 1897 houses, addresses were used to match their structural characteristics with their most recent sale price recorded by the county Assessor⁴⁵.

The next step was to assign each house to its respective community. First, GIS software was used to translate the address of each house into latitude/longitude coordinates. Then, the coordinate points were overlaid on the map of communities to make the assignment. This process was done using the geocoding application built into *ArcView GIS 3.2*, along with *Census 2000 Tiger* street maps. *ArcView*’s geocoding algorithm assumes houses are equidistantly spaced on each street. The error introduced by this interpolation is

⁴⁵ I assume that the houses currently listed for sale on *Metrolist* have not had major structural additions since their last sale date.

of little concern since it should not affect how houses are assigned to communities. A more important issue is that 22% of the houses could not be geocoded, reducing the sample size to 2.2 million observations. The primary reason is that the Tiger maps are incomplete and tend to lag behind current development⁴⁶. As a result, newer homes may be disproportionately underrepresented. This could actually mitigate concerns that newer homes are overrepresented in the data because every new house sold during the study period will be included in the sample, whereas older homes will only be included if they happen to be resold during the study period.

The geocoded data were filtered to remove observations missing one or more of the core structural characteristics, apparent errors (e.g. houses with 128 bathrooms), nonresidential properties (e.g. commercial buildings, cemeteries, etc), and outliers—specifically the most expensive and least expensive 0.5% of sales. Houses sold before 1995 were also eliminated out of the concern that the supply of public goods may have changed over time. Finally, prices were converted to year 2000 dollars using the California consumer price index. The filtering process reduced the sample size to 540,624 transactions between 1995 and 2005—the final data used to estimate the housing price index. Table 4.3 summarizes these observations. Notice that the average sale year is 2000, suggesting that, on average, newer homes are not over-sampled. In addition to the characteristics shown in the table, the sale price of a house is also likely to be influenced by whether it has a pool, fireplace, garage, or central heating/air to name just a few. Unfortunately these attributes are not reported by the assessors in roughly half the counties, nor are they included in the set of characteristics reported by the online sources used for Yolo County. If these attributes differ systematically across communities, their marginal implicit price should be incorporated into the unobserved public good recovered as part of the estimation. Otherwise, they will simply introduce noise into the hedonic regression used to calculate the housing price index.

In addition to calculating the price index, the data on individual housing transactions were used to estimate the demand for housing. Because the estimating equation (3.10)

⁴⁶ ArcView adjusts for obvious variations in spelling like the difference between “Zephyr Ct” and “Zephyr Court”, and also permits a small share of spelling errors. Egregious spelling errors may have prevented some houses from being geocoded, although the effect was small in comparison to the issue of street names that were missing in the *Tiger* street files.

defines housing expenditures as an annualized measure, the 540,624 observations on price were converted into rents using the formula suggested by Poterba (1992). Equation (4.1) shows the relationship between the sale price (P) and the annualized rental price (R) of a home, and table 4.4 defines the variables in the equation along with the values that were assigned to each.

$$(4.1) \quad R = [(1 - \tau)(i - \tau_p) + r + m + \delta - \pi]P .$$

The marginal tax rate (τ) of 23% is the average marginal tax rate for households in the U.S. as reported by Saez (2004); California's Proposition 13 fixes the annual property tax rate (τ_p) at 1%; and the 7% interest rate (i) represents an annual average of the 30-year fixed rate mortgage as reported by the Federal Home Loan Mortgage Corporation from 1995 to 2005. The annual risk premium ($r = 4\%$), maintenance rate ($m = 2\%$), and depreciation rate ($\delta = 2\%$) were all set to Poterba's suggested values. Finally, the land appreciation rate ($\pi = 5\%$) was calculated as the annual average inflation rate for the consumer price index of housing in the San Francisco CMSA from 1995 to 2005, as reported by the Bureau of Labor Statistics⁴⁷. The resulting formula implies the annual rental rate for housing equals 9.17% of its price.

(B) *Air Quality: Ozone*

Ozone concentrations are used as a proxy for air quality. Ozone is an attractive proxy because it is documented to have negative human health effects, particularly on respiratory tract tissue, and is the chief component of urban smog, making it readily observable. Moreover, local news stations and the major regional newspapers (the *Sacramento Bee* and *San Francisco Chronicle*) report EPA's Air Quality Index (AQI) in their weather section. The AQI is based on a composite of ozone and other highly correlated pollutants such as particulate matter. It condenses the index of air quality into color-coded indicators such as "good", "unhealthy for sensitive groups", and "very unhealthy". Ozone has also been found

⁴⁷ A comparable measure for the Sacramento CMSA was not available. It seems likely that the San Francisco measure would be more representative of the Sacramento CMSA than a statewide measure.

to affect housing prices in empirical hedonic and sorting studies.

The California Air Resources Board records hourly concentrations of ozone at monitoring stations throughout the state. Figure 4.2D overlays the location of 210 monitoring stations on school districts in the study region. The ozone measure used in this analysis is the average of the top 30 1-hour daily maximum readings (in parts per million) recorded at each monitoring station during the course of a year. This is the same measure used in the locational equilibrium studies by Sieg et al. (2004) and Smith et al. (2004). While alternative air pollution measures are available, the fact that negative health effects are triggered at higher ozone concentrations suggests that the highest concentrations recorded at each station may be the most relevant.

Community-specific measures of air quality are constructed by first assigning to each house the ozone measure recorded at the nearest monitoring station, and then taking an average over all the houses in the community. Then, to control for annual weather-related fluctuation in ozone levels, the process was repeated for 1999, 2000, and 2001, and the results averaged. The final measure ranges from 0.031 in the highest air quality community to 0.106 in the lowest. Figure 4.3A shows the spatial distribution of air quality within the region. Not surprisingly, ozone concentrations are lowest near the coast and then increase as one moves further inland, with the unhealthiest air directly east of the city of Sacramento. While there is substantial variation in air quality within each CMSA, ozone levels are systematically higher in the Sacramento region. Part (D) discusses the implications of this spatial correlation.

(C) *School Quality: Academic Performance Index*

Because the *Serrano vs. Priest* (1971) decision by the state Supreme Court effectively equalized expenditures per student across school districts, public school quality in California is perhaps best measured by academic performance. The measure used in this study is the Academic Performance Index (API), which was created by the California Public Schools Accountability Act of 1999 to be an objective measure that could be used by legislators and parents to compare the state's public schools. It is a composite index of standardized test

scores, weighted across all subjects and grade levels. Since January, 2000, schools have been required to annually report their current API score to parents, and to explain the ranking system. Parents considering relocating to a different district can obtain API scores of individual schools directly from the California Department of Education, or on its website.

A three-year average API was constructed for each school district by weighting the score of each school in that community by its student population during 1999-2001. The resulting measure ranges from 528 to 941. Figure 4.3B shows the spatial distribution of school quality in the region. Ranked by API, all but one of the top 20% of communities are located around the San Francisco Bay. The lone exception is the Davis Unified school district in Yolo County, part of the Sacramento CMSA.

(D) *Unobserved Public Goods*

While school quality and air quality are the primary focus of this study, other local public goods undoubtedly influence location choices. Examples include crime, microclimate, dimensions of “open space”, and proximity to entertainment, dining, and other cultural amenities. Recall that the composite unobserved public good that is recovered during the estimation can be thought of as an index of these and other unobserved amenities. For example, if this index were assumed to be linear, it could be represented by (4.2) where the ω ’s are relative weights on each of the M unobserved public goods provided by community j .

$$(4.2) \quad \bar{\xi}_j = \omega_1 \xi_{j,1} + \omega_2 \xi_{j,2} + \dots + \omega_M \xi_{j,M}.$$

For this index to be a theoretically consistent measure, households must have vertically differentiated preferences for each unobserved public good; i.e. the weights must be constants⁴⁸. This is true regardless of whether the index is linear or nonlinear.

⁴⁸ This restriction should not be confused with the depiction of preference heterogeneity for the observed public goods. The specification for utility in equation (3.1) depicts households that have horizontally differentiated preferences for the observed public goods.

The key assumption that allows an index of unobserved public goods to be recovered during the estimation is that the index is uncorrelated with air and school quality, as defined in equation (3.11.b). Since air and school quality both tend to be higher in the San Francisco region, one might worry that other public goods with a similar spatial distribution would cause the independence assumption to be violated. When faced with a similar pattern of air quality in the Los Angeles metropolitan area, Sieg et al. (2004) included *distance from the coast* as an additional characteristic in their hedonic regression, removing that effect from the index of housing prices. The idea is that controlling for distance to the coast would remove the effect of beach access as well as other public goods that share a similar spatial distribution. A different strategy would be to include *distance from the coast* as an additional community-specific public good. A disadvantage of that approach is it would increase the computational burden by expanding the dimensionality of the Gibbs sampling process. Alternatively, one could argue that even if some unobserved public goods are correlated with observed ones, the index as a whole could still be independent. For example, suppose that one unobserved public good ($\xi_{j,1}$) in the index in (4.2) is correlated with air quality. As long as ω_1 is sufficiently small, the independence assumption should still approximately hold.

(E) *Commute Time*

The San Francisco-Sacramento region is one of the most congested in the United States. Ranked by average commute time, eight of its nine PMSAs are among the top 60 out of 338 in the U.S. The exception—Yolo again—is ranked at #214. Working households in the region are assumed to value the time they spend commuting and to make location choices that reflect the opportunity cost of their time. A worker who values her time highly may choose a job that pays less if it requires a shorter commute. Likewise, she may choose to live in a community with fewer public goods if it is close to a higher paying job. To characterize this set of potential tradeoffs, commute times must be attached to each of the 268 (community, work destination) pairs.

Data on commuting time were taken from the Census Transportation Planning Package special tabulation, which gives the mean time for every tract-to-tract commute

reported by workers in the 2000 census. The data are available by mode of transport (car, bus, bike, etc.) and by time of day. These figures were aggregated to generate a weighted average travel time between each home community and PMSA. The weights were defined by the number of workers observed making each (mode of transport, time of day, tract-to-tract) commute. The resulting average one-way commute time ranges from 1 to 114 minutes, with a mean of 36 minutes and a standard deviation of 19 minutes. Traffic is a major contributor to the relatively high average commute time. Most workers (82%) live and work in the same PMSA (table 4.2).

(F) *Wages*

A worker's job opportunities depend on his skills. In the empirical model, the set of labor market choices faced by each working household is defined by the spatial distribution of wages that corresponds to that household's occupation. Each working household is classified according to the occupational category of its primary earner, using the Standard Occupational Classification System developed by the Bureau of Labor Statistics. For each of these 22 categories, the distribution of wages across Primary Metropolitan Statistical Areas was obtained from the California Employment Development Department (EDD).

EDD surveys roughly 37,000 non-farm establishments each year, asking them to report the wage paid to each worker, and to classify each according to the 821 job titles defined by the Standard Occupational Classification System⁴⁹. Specific job titles include *short order cook*, *fire fighter*, and *economist*. If employers report wages on an hourly basis, EDD converts these figures to an annual basis by assuming 52 40-hour work weeks, adjusting for jobs where workers are paid for fewer than 12 months, such as teachers. "Wages" include base pay, production bonuses, tips, and cost-of-living adjustments, but exclude nonproduction bonuses, overtime pay and the value of benefits. PMSAs are the smallest geographic unit for which these data are publicly available.

⁴⁹ EDD collects these data as part of the Occupational Employment Statistics Survey, developed by the Bureau of Labor Statistics.

The mean annual wage for each occupational category is an average over the wages for all the specific job titles in that category, weighted by the number of workers. Mean annual wages were then averaged over 1999, 2000, and 2001 to generate the wage data used in the estimation. The resulting average wage rates can vary substantially between PMSAs, as table 4.5 demonstrates. Workers with jobs in the *construction and excavation* category are paid 32% more in San Jose than in Sacramento, for example. Some of this variation may reflect local cost-of-living adjustments in markets where housing is particularly expensive, like San Jose and San Francisco. The variation may also reflect location-specific attributes of certain jobs or heterogeneity in the mix of jobs within each category.

Aggregating over job titles misspecifies the actual labor market options faced by working households. One aspect of this problem is illustrated by Table 4.6, which shows how dramatically wages can vary within the occupational categories. The empirical model mitigates this problem somewhat by allowing workers to differ in an unobserved component of their job skill (θ_1). Suppose we observe *protective service* workers living in the Vallejo area and commuting to Oakland. This choice is difficult to justify based on the spatial distribution of wages. Table 4.5 shows they could earn more by working in the Vallejo PMSA, presumably a shorter commute. Yet commuting from Vallejo to Oakland would be an understandable choice for the average *fire fighter* who would be paid \$13,500 more in Oakland, as can be seen from table 4.6. The estimator reconciles the apparent inconsistency in the Vallejo-to-Oakland commute by assigning a value of θ_1 that recognizes a *protective service* worker who chose to work in Oakland must have done so because their particular skill set qualifies them for higher pay there than in Vallejo. The higher wage in Oakland effectively decreases the firefighter's labor market mobility. Other factors that could affect labor market mobility include idiosyncratic characteristics of the worker and location-specific attributes of the job. While the flexibility introduced by θ_1 helps to mitigate aggregation problems, it does not eliminate them. If aggregating over job titles misspecifies a working household's labor market options, it will lead to biased estimates for their preferences and job skill.

Households with more than one worker pose another potential source of bias. Recall that the empirical model treats wages as fixed for workers other than the primary earner of a household. That is, wages paid to all non-primary earners are included in exogenous non-wage income. If these workers also face a variety of wage-commute options, the empirical model will misrepresent their household's labor market choices. Data on multi-earner households suggests this could be an important problem. Table 4.7 shows the number of households with 1, 2, and more than 2 workers by PMSA. Just over half of the 2.5 million working households in the study region have more than one worker. By PMSA, the share of households with more than one worker ranges from 48.8% in Sacramento to 54.4% in San Jose. However, these shares will overstate the size of the problem to the extent that there are households where the secondary earner's income is incidental or effectively exogenous. For example, the secondary earner could be a high school student with an after-school job, an adult that telecommutes, or someone who is self-employed.

In theory, the empirical model could be extended to allow multiple-earner households and it could be estimated using data on the job opportunities available to workers with each of the 821 individual job titles. This would reduce concerns about aggregating over diverse skill sets and treating wages as fixed for a large share of workers. Unfortunately, the publicly available data only report occupations and job locations for the primary earner of a household. Likewise, the 22 occupational categories are the most detailed level at which location choices can be observed. Section V discusses these limitations in greater detail.

V. Observing Location Choices by Income and Occupation

Having defined the choice set, the last step in preparing the data is to observe choices. The challenge is that location choices must be observed together with households' income and occupations. This is because the partitioning stage of the estimation is conditional on both non-wage income and occupation. (Recall that because utility depends on both non-wage income and occupation, so does the partition of preference space.) It is relatively straightforward to observe community-level income distributions for retired households.

Unfortunately, the same is not true for working households. Location choices, income, and occupations cannot be observed simultaneously in publicly available data. After briefly describing how the data were generated for retired households, the rest of this section describes how multiple data sources were combined to infer incomes, occupations, and location choices for working households.

The distribution of income for retired households in each of the 122 communities was inferred directly from the year 2000 Census of Population and Housing. The special tabulation of these data provided by the Census Transportation Planning Package (CTPP) reports the distribution of income by census tract for *households with no workers*. These tract-level income distributions were aggregated into housing communities under the assumption that households are distributed uniformly (in space) across each tract and that “households with no workers” are retired⁵⁰. A disadvantage of this approach is that it will probably misclassify some temporarily unemployed households as being retired. Fortunately, the magnitude of the problem is likely to be small since the unemployment rate in the study region was only 3% in the year 2000 (CA Statistical Abstract). An alternative way to identify retired households would be to use income distributions by age, making an assumption about the threshold age for retirement. That approach, however, seems likely to introduce a greater margin of error given the variability in retirement age. For example, in the year 2000, 25% of Americans aged 65-69 were in the labor force, as were 13.5% of those aged 70-74 and 5.3% of those over 75 (Wiatrowski [2001]).

For working households, the task is more complex. The distribution of income must be collected for workers in each of the 268 job-house locations, by occupational category⁵¹. Table 4.8 provides an example of what the required data should look like. The Census Bureau collects this information but, unfortunately, does not release it to the general public in one piece. Instead, the Bureau reports the individual components of the table in separate

⁵⁰ The aggregation was done using ArcView GIS 3.2. Technically, the uniformity assumption is only required in situations where census tracts overlap multiple communities, which occurs rarely.

⁵¹ Non-wage income is calculated by subtracting the primary earner’s wage income from total household income. As discussed earlier, this treats the wage income for all other earners in the household as being exogenous to location choice.

special tabulations of Census 2000 data: the CTPP and the Equal Employment Opportunity Package (EEO). The strategy developed here is to combine these two special tabulations with the wage data described in the previous section to generate an approximation to the required data. More specifically, I make a series of assumptions about the conditional distributions of *home tract by occupation*, *commuting pattern by occupation*, and *household income* in order to generate an approximation to their joint distribution. Table 4.9 summarizes the key assumptions and the specific variables that were used. The data-generating process is described in three steps.

The first step is to combine the two special tabulations to determine commuting patterns for workers in each occupational category. CTPP reports the number of workers in each occupation, by home census tract. EEO reports county-to-county commuting patterns by occupation⁵². To infer tract-to-county commuting patterns, I assume county-level commuting is representative of tract-level commuting. The assumption is formally expressed as (4.3) where $S_{Y,t \rightarrow X}$ represents the number of workers in occupation S who live in tract t of county Y and commute to county X , and $S_{Y,t}$ represents the total number of occupation S workers living in the tract. $S_{Y \rightarrow X}$ and S_Y simply aggregate over all the tracts in county Y .

$$(4.3) \quad \frac{S_{Y,t \rightarrow X}}{S_{Y,t}} = \frac{S_{Y \rightarrow X}}{S_Y} \quad \forall S, Y, X, t.$$

Multiplying both sides of the equation by $S_{Y,t}$ places all the observable variables to the right of the equality and allows me to calculate the number of workers commuting between each home tract and work county. These results were used to generate the number of workers commuting between each home tract and work PMSA using the fact that PMSAs are a direct aggregation of counties.

The second step in generating the data is to combine the commuting data with information on wages and household income. It is straightforward to attach the mean wage

⁵² For each destination county, the number of workers commuting to that county is only reported for the top 9 origin counties. However, given the commuting patterns in table 4.2, it seems likely the top 9 counties will capture virtually all workers.

for each occupational category to the data on commuting patterns. Table 4.10 provides an example of the result. To combine this data with information on total household income requires the following key assumption: *within each census tract the distribution of total household income, conditional on wages, is independent of the worker's occupational category*. This is stated as (4.4).

$$(4.4) \quad f(y_i) | w_1 = f(y_i) | w_2 = \dots = f(y_i) | w_S \quad \text{if} \quad w_1 = w_2 = \dots = w_S.$$

The assumption implies that career choice itself does not affect other sources of household income like a spouse's wages and return on investments. This allows the wage data in table 4.10 to be combined with tract-level distributions of total household income by worker earnings, which are aggregated over all occupations⁵³. To see how this works, consider the 271 managers in the table above who live in tract 400100 and work in the Oakland PMSA. Their mean wage is \$78.2k. Table 4.11 displays the distribution of household income for all workers in the tract with earnings over \$75k reported by the CTPP tabulation. The percentages in the bottom row of the table are simply multiplied by 271 to get the distribution of household income for managers: 10 with income from \$75-\$100, 22 with income from \$100-\$125, and 239 with income over \$125. Table 4.12A shows results from repeating this logic for all workers in the tract.

Notice that, in the summary lines at the bottom of Table 4.12A, the distribution of household income observed in the CTPP tabulation contains 922 households whereas the distribution generated by the process described in the previous paragraph contains 1470 households. The discrepancy is caused by households with more than one worker. That is, the generated distribution is based on the universe of *workers in households* whereas the CTPP data are based on the universe of *working households*. The discrepancy is a pervasive

⁵³ Two additional technical assumptions are required to guarantee compatibility between the EDD definition for *wages* and the Census definition of *earnings*: (1) Self-employed workers are paid the same as salary workers; (2) Each worker holds exactly one full-time job where they work 40 hours a week. An additional minor concern stems from the fact that the EDD data come from the universe of *workers*, while the CTPP data come from the universe of *workers in households*. The former includes workers living in group quarters (e.g. military housing, college dorms); the later does not. The concern is that it seems unlikely that workers living in group quarters have the same distribution of occupations as those who do not. However, only 2.5% of census tracts have more than 5% of workers living in group quarters, which is why the concern seems minor.

problem; it affects every census tract in the study region. Overall, the region contains 4.2 million workers and only 2.5 million working households. Likewise, 60% of married couples reported both the husband and wife working in 1999.

The final step of the data-generating process converts the income distribution generated for all workers to a distribution based on the primary earners in working households. This is consistent with the maintained assumption that each working household makes its location choice based on the job options faced by its primary earner. A primary earner is defined here as the worker who accounts for the largest share of a household's wage income. All other earners are removed from the distribution by assuming that, conditional on household income in a census tract, the maximum wage of any household member who is not a primary earner is less than or equal to the minimum wage of any primary earner. Equation (4.5) states this condition formally, where the *NPE* and *PE* superscripts indicate that the individual *is not* or *is* the primary earner of a household.

$$(4.5) \quad \max(w^{NPE})|y \leq \max(w^{PE})|y.$$

If (4.5) holds, the income distribution for primary earners can be recovered by solving for the difference between the total number of workers in the generated and observed income distributions in Table 4.12A (call it *D*), and then removing the *D* workers in each generated income bin with the lowest wages. For example, the generated distribution reports 894 households with income over \$125k while CTPP only reports 550. Therefore the 344 households with the lowest wages in that income bin are removed. The idea is that these workers are less likely to be primary earners of their household because they account for a smaller share of total household income. Table 4.12B shows the resulting income distribution.

The totals at the bottom of Table 4.12B still do not match. The reason is that the generated distribution in Table 4.12A understates the number of workers in three of the income bins. One explanation for this is that some primary earners only work part time so assigning them an annual wage based on a 40-hour work week greatly overestimates their wage income, causing them to be placed in the wrong earnings bin. This is consistent with

the observation that the generated income distribution tends to underestimate the number of workers in the lower income bins⁵⁴. Consequently, these workers are treated as being retired and are added to the distribution of retired households. The rationale is that a worker who retires from her career and then chooses to work a part-time job in order to supplement retirement income, or just for fun, is nearly as mobile as a retired household with no job. That is, she could find some kind of part-time work anywhere. The drawback of this approach is that it may incorrectly place some full-time workers with wages far below the mean for their occupational category in the retired income distribution.

Finally, *ArcView* GIS software was used to aggregate the resulting income distributions from the 1808 home Census tracts into the 122 communities. Combining this information on choices with the choice set developed in sections III and IV gives the final data set used in the estimation. Table 4.13 provides a sample of these data.

VI. Conclusions

This chapter developed the data required to estimate the empirical model. First, the study region was defined as the union of the San Francisco and Sacramento Consolidated Metropolitan Statistical Areas. Recent migration patterns were used to define the geographic boundaries of this region, which was subsequently subdivided into 122 housing communities and 8 work destinations. Then commuting patterns were used to reduce the set of all possible job-house combinations to 268 locations. For each location, data were collected on proxy measures for two public goods (ozone for air quality and the Academic Performance Index for school quality), on a set of wage rates paid to workers in each of 22 occupational categories, and on commute times. The final step was to collect data on household income, occupations, and location choices. Because their joint distribution could not be observed directly, it was approximated by combining conditional distributions available in public data sets. Table 4.14 shows descriptive statistics for the 122 housing communities. Ozone levels

⁵⁴ Another explanation is that workers in the census tract are paid much more or much less than the mean for their occupational category so that assigning them the mean wage places them in the wrong earnings bin. This problem would be difficult to address without additional information.

and school quality vary substantially. Likewise, quantiles of the distribution of household income and housing expenditures differ by up to an order of magnitude between communities.

The size of the San Francisco-Sacramento area is typical of sorting applications. Epple and Sieg (1999) use the Boston MSA, Sieg et al. (2004) and Smith et al. (2004) both use the Los Angeles CMSA, and Bayer et al (2005) use a subset of the San Francisco CMSA. It seems the concept of a metropolitan area (or in this case two adjacent ones) is capable of satisfying the restrictions on mobility that drive the revealed preference logic of the model, as well as satisfying practical limits on data and computing power. The idea that a metropolitan area (or some subset thereof) can be treated as a distinct market is also regularly used in reduced-form hedonic property value studies (Palmquist [2005]). However, there are other hedonic studies that define the choice set as the entire United States, treating each MSA as a single observation and ignoring moving costs entirely (e.g. Roback [1982] and Chay and Greenstone [2005]). Bayer et al. (2006) find that incorporating moving costs into such a model triples their estimates for the marginal willingness-to-pay for air quality. The estimator in chapter 3 could be extended to incorporate moving costs if data on the number of recent movers were also available. Though it was argued that the cost of moving within the San Francisco-Sacramento region is relatively small, it is certainly not zero.

Including job locations as a dimension of choice complicated the data collection process relative to previous sorting studies that focused solely on house locations (Epple-Sieg [1999] and Sieg et al. [2004]) or that measured commuting time but treated job locations as fixed (Bayer et al. [2005]). The way a choice was defined also differentiates the structure of the choice set from interregional hedonic models which implicitly confine a household's job and house to the same metropolitan area (Roback [1982], Blomquist, Berger, and Hoehn [1988]). The set of 268 job-house locations recognizes that in a region with multiple metropolitan areas workers may choose to live in one MSA and work in another.

An important caveat to the final data set is that numerous assumptions were required to approximate income distributions by location and occupation. An alternative to making these assumptions would be to obtain access to the restricted census micro data available

only at Census data centers. For example, access to the census micro data would make it possible to relax the assumption that households sort on the basis of job opportunities faced by their primary earner. The estimator treats wages of secondary earners as a component of exogenous non-wage income. Yet a household with two earners may have to find new jobs for both if it moves sufficiently far from its current location. This type of constraint on mobility may apply to a large number of households—approximately half the working households in the study region have more than one wage earner. Chapter 7 returns to proposal to use Census micro data to model a dual-earner job search as a possibility for future research.

Table 4.1: Land Use, Employment, and Migration in Northern California MSAs, 1995-2000

Metropolitan Area	Share of Land in Farms	# Employees (1,000)		# People Moving From:	
		Total	Agricultural	Sacramento	San Francisco
San Francisco	45%	3,661	36	41,185	483,928
Sacramento	34%	806	9	72,070	87,369
Stockton	91%	203	17	5,902	41,679
Modesto	83%	160	16	2,081	23,342
Salinas	59%	166	39	2,143	20,822
Merced	82%	64	12	946	10,910
Yuba	77%	43	6	4,256	2,380

Sources: (1) California Statistical Abstract 2005, California Department of Finance.
(2) California Employment Development Department, Labor Market Information Division.
(3) County-to-County outflow files, Census of Population, 2000.

Table 4.2: Work Location, by Home PMSA, 2000*
(percentage of total workforce)

Home PMSA	Work PMSA								
	Sacramento	Yolo	Oakland	San Francisco	San Jose	Santa Cruz	Santa Rosa	Vallejo	Other
Sacramento	90.6	3.7	0.6	0.5	0.3	0.0	0.1	0.6	3.6
Yolo	24.5	67.2	1.1	0.9	0.2	0.0	0.1	4.3	1.7
Oakland	0.2	0.0	74.9	15.6	7.1	0.1	0.2	0.9	1.0
San Francisco	0.1	0.0	5.5	84.8	8.0	0.1	0.5	0.2	0.8
San Jose	0.0	0.0	4.8	5.9	87.8	0.4	0.1	0.1	0.8
Santa Cruz	0.0	0.0	1.3	2.2	17.1	73.8	0.1	0.1	5.3
Santa Rosa	0.1	0.0	1.8	12.5	0.6	0.0	82.0	1.9	1.0
Vallejo	2.2	1.6	16.3	8.8	0.8	0.0	1.9	67.1	1.3

* For example, 24.5% of Yolo's working population works in the Sacramento PMSA.

Source: County-to-County Worker Flow Files, Census of Population, 2000.

Table 4.3: Summary Statistics for Housing Sales, 1995-2005 (N=540,624)

Characteristic	Mean	St. Dev.	Min	Max
sale price (year 2000 dollars)	373,106	247,642	52,363	1,899,000
year sold	2000	2.80	1995	2005
lot size (sqft)	12,910	39,662	435	871,200
structure (sqft)	1,667	682	400	149,608
age	32	23	1	154
# bedrooms	3.20	0.84	1	15
# bathrooms	2.05	0.70	0	15

Source: *Dataquick*.

Table 4.4: Variables Used to Convert Housing Prices into an Annualized Rental Price

Variable	Symbol	Value	Source
owner's marginal tax rate	τ	23.16%	Saez (2004)
property tax rate	τ_p	1.00%	Prop. 13
interest rate	i	6.98%	Freddie Mac
risk premium	r	4.00%	Poterba (1992)
maintenance	m	2.00%	Poterba (1992)
depreciation	δ	2.00%	Poterba (1992)
land appreciation rate	π	4.96%	BLS

Table 4.5: Mean Annual Wages (\$1000) by Occupational Category, 1999-2001

Occupational Category (Standard Occupational Classification System)	SOC Code	Primary Metropolitan Statistical Area							
		San Jose	Oakland	Sacramento	San Francisco	Santa Cruz	Santa Rosa	Vallejo-Fairfield-Napa	Yolo
Management	110	91.1	78.2	68.6	82.4	73.1	70.6	70.8	66.0
Business and Financial Operations	130	55.9	53.4	44.6	58.2	47.2	46.8	49.7	46.3
Computer & Mathematical	150	71.8	64.6	57.6	67.1	64.0	58.5	49.3	48.4
Architecture and Engineering	170	64.4	57.7	54.5	58.5	54.9	54.3	54.9	50.1
Life, Physical and Social Science	190	58.4	55.8	47.4	54.2	39.3	49.4	46.6	46.8
Community and Social Service	210	36.2	34.2	33.6	34.1	32.4	29.7	37.5	26.7
Legal	230	83.8	62.4	73.3	82.7	64.3	49.6	54.5	62.0
Education, Training and Library	250	41.2	39.8	38.1	40.1	38.8	39.0	37.5	36.1
Arts, Design, Entertainment, Sports and Media	270	50.3	44.0	39.5	49.0	39.8	37.5	35.1	32.1
Healthcare Practitioners and Technicians	290	56.7	55.5	48.9	55.3	53.8	45.5	47.2	50.6
Healthcare Support	310	27.4	25.4	22.8	25.3	25.2	25.0	24.7	21.9
Protective Service	330	37.0	38.3	34.1	37.8	39.6	34.8	40.1	32.4
Food Preparation and Serving	350	16.6	16.8	16.0	18.3	16.6	16.6	16.3	14.5
Building and Grounds Cleaning & Maintenance	370	20.9	23.2	20.4	22.9	20.5	20.2	20.1	17.9
Personal Care and Service	390	26.7	22.1	19.5	27.7	21.2	21.3	21.2	19.4
Sales and Related	410	35.5	31.3	27.2	37.9	25.5	27.3	25.1	26.2
Office and Administrative Support	430	32.5	29.5	27.1	32.2	26.6	27.6	26.0	25.5
Farming, Fishing and Forestry	450	17.3	24.5	20.4	24.5	16.5	19.1	18.6	18.3
Construction and Excavation	470	42.5	43.6	34.6	45.6	37.3	37.7	42.1	37.7
Installation, Maintenance and Repair	490	36.8	36.0	33.4	37.2	34.2	35.0	36.7	33.6
Production	510	28.9	29.3	24.6	28.0	25.4	25.8	27.4	23.4
Transportation and Material Moving	530	25.0	27.7	23.9	28.7	23.3	24.2	25.0	26.4

Source: California Employment Development Department, Occupational Employment Survey, 1999-2001.

Table 4.6: Mean Annual Wage for Selected Occupations by PMSA, 2001-2002

SOC Category (22)	Office and Administrative		Legal		Protective Service		Healthcare Practitioners	
	File Clerks	Legal Secretaries	Lawyers	Paralegals and Legal Assistants	Fire Fighters	Security Guards	Pharmacists	Registered Nurses
Oakland	26,768	54,404	114,234	53,005	69,193	23,431	86,424	73,588
Yolo	22,472	31,437	110,235	44,012	40,308	26,591	88,002	55,853
Sacramento	22,541	42,150	103,753	43,151	41,500	20,008	84,426	58,913
San Francisco	27,338	55,385	149,402	54,105	59,477	24,591	82,305	70,790
Santa Rosa	22,322	35,627	102,161	27,821	31,000	22,525	85,236	58,876
Vallejo	24,490	36,153	94,166	36,935	55,685	26,948	81,793	58,515
San Jose	28,854	52,929	154,406	54,944	61,998	25,963	85,404	68,344
Santa Cruz	21,332	43,314	98,091	39,518	48,010	20,936	86,258	58,268

Source: California Employment Development Department, Occupational Employment Survey, 1999-2001.

Table 4.7: Number of Single, Dual, and Multi-Earner Households, by PMSA (thousands)

Home PMSA	All Working Households	Household Type			Multi-Earner Households (share)
		Single Earner	Dual Earner	More than 2 Earners	
Sacramento	455	233	184	38	0.49
Yolo	45	22	19	4	0.51
Oakland	677	331	272	73	0.51
San Francisco	534	264	205	65	0.51
San Jose	473	216	188	69	0.54
Santa Cruz	73	34	29	9	0.53
Santa Rosa	132	62	55	15	0.53
Vallejo	136	63	57	16	0.54
Total	2,524	1,225	1,010	289	0.51

Source: Census Transportation Planning Package, 2000.

Table 4.8: An Example of the Required Data

Occupation of Primary Earner	Home Community	Work PMSA	# Households with Total Household Income (\$1000)									
			<10	10-15	15-30	30-40	40-50	50-60	60-75	75-100	100-125	>125
Management	Oakland Unified	Oakland	0	0	0	0	0	0	0	123	59	174
Management	Oakland Unified	San Jose	0	0	0	0	0	0	0	74	21	59
Management	Oakland Unified	Vallejo	0	0	0	0	0	0	0	37	6	1

Table 4.9: Data and Assumptions Used in the Data-Generating Process

Variable*	Years	Spatial Aggregation	Universe	Source
Mean annual wage, by occupation (821)	1999-2001	PMSA	workers	California Employment Development Department
# households, by # workers in household (5) & household income (25)	2000	tract	households	CTTP Special Tabulation
# workers, by occupational category (22)	2000	tract	workers	CTTP Special Tabulation
# households by income (10) & worker earnings (10)	2000	tract	workers in households	CTTP Special Tabulation
# workers commuting, by occupation (590)	2000	county to county	workers from top 9 origin counties	EEO Special Tabulation

* The numbers in parentheses refer to the number of unique values for that variable. For example, "# workers, by occupational category (22)" refers to the fact that # workers is reported for each of 22 specific occupational categories. The 22 occupational categories are an exact aggregation of the 590 reported in the EEO special tabulation.

Assumptions Used in the Data-Generating Process	
A1	For each census tract t in county Y : if any workers from tract t report commuting to county X , then the share of workers making the commute <u>for each occupational category</u> in tract t is the same as the share of workers in that <u>occupational category</u> that are reported as residing in county Y and making the commute from county Y to county X .
A2	Self-employed workers are paid the same as salary workers for each occupation.
A3	The primary earner of each household holds exactly one full-time job where they work 40 hours a week.
A4	The distribution of household income, conditional on wages of the primary earner, is independent of the primary earner's occupational category.
A5	For each census tract, conditional on household income, the maximum wage of any household member who is not a primary earner is less than or equal to the minimum wage of any household member who is a primary earner.

Table 4.10: Combining Data on Commuting Patterns by Occupation with Data on Wage Income, an Example

Home: Tract 400100, Oakland Unified School District			
Occupation	Work PMSA	Wage (\$1000)	# Workers
Management	Oakland	78.2	271
Management	San Jose	91.1	66
Management	Vallejo	70.8	1

Table 4.11: An Example of the Conditional Income Distributions Reported in the CTPP Special Tabulation

Tract 400100: Distribution of Household Income for Workers with Earnings > \$75,000								
HH Income	<30	30-40	40-50	50-60	60-75	75-100	100-125	>125
# Workers	0	0	0	0	0	25	55	600
% Workers	0.0%	0.0%	0.0%	0.0%	0.0%	3.7%	8.1%	88.2%

Table 4.12A: Generating Income Distributions by Occupation and Location

Wages and Commuting Patterns				Household Income (\$1000)								Primary Earners
Occupation	Work PMSA	# Workers	Wage	<30	30-40	40-50	50-60	60-75	75-100	100-125	>125	
Management	Oakland	271	78.2	0	0	0	0	0	10	22	239	271
Management	San Francisco	83	82.4	0	0	0	0	0	3	7	73	83
Management	San Jose	66	91.1	0	0	0	0	0	2	5	59	66
Management	Vallejo	1	70.8	0	0	0	0	0	0	0	1	1
Business & Financial	Oakland	102	53.4	0	0	2	2	7	7	15	69	102
Business & Financial	San Francisco	39	58.2	0	0	1	1	3	3	6	25	39
Business & Financial	San Jose	18	55.9	0	0	0	0	1	1	3	13	18
.
.
.
Generated Income Distribution for Working Households:				6	84	12	17	91	154	212	894	1470
Observed Income Distribution for Working Households:				43	62	25	32	45	75	90	550	922

Sources: (1) Census Bureau, Census Transportation Planning Package Special Tabulation, 2000.
 (2) Census Bureau, Census Equal Employment Opportunity Special Tabulation, 2000.
 (3) California Employment Development Department, Occupational Employment Survey, 1999-2001.

Table 4.12B: Converting the Universe of Workers in Households to Working Households

Wages and Commuting Patterns				Household Income (\$1000)								Primary Earners
Occupation	Work PMSA	# Workers	Wage	<30	30-40	40-50	50-60	60-75	75-100	100-125	>125	
Management	Oakland	271	78.2	0	0	0	0	0	10	22	239	271
Management	San Francisco	83	82.4	0	0	0	0	0	3	7	73	83
Management	San Jose	66	91.1	0	0	0	0	0	2	5	59	66
Management	Vallejo	1	70.8	0	0	0	0	0	0	0	1	1
Business & Financial	Oakland	102	53.4	0	0	2	2	8	8	0	0	20
Business & Financial	San Francisco	39	58.2	0	0	1	1	3	3	5	16	29
Business & Financial	San Jose	18	55.9	0	0	0	0	1	1	3	0	5
.
.
.
Final Income Distribution for Working Households:				6	62	12	17	45	75	90	550	857
Observed Income Distribution for Working Households:				43	62	25	32	45	75	90	550	922
Additional "Retired" Households:				37	0	13	15	0	0	0	0	65

Sources: (1) Census Bureau, Census Transportation Planning Package Special Tabulation, 2000.
 (2) Census Bureau, Census Equal Employment Opportunity Special Tabulation, 2000.
 (3) California Employment Development Department, Occupational Employment Survey, 1999-2001.

Table 4.13: Sample of the Data used in Estimation: Sales & Related Occupations (SOC 410)

Occupation & Location			Public Goods		Job Attributes		Constructed Household Income Bins (\$1000)									
Occupation	Home Community	Work PMSA	School	Air	Commute time share	Wage income	<10	10-15	15-30	30-40	40-50	50-60	60-75	75-100	100-125	>125
Sales	Acalanes	Oakland	1.20	0.74	0.10	31,306	0	0	0	957	244	142	83	54	31	32
Sales	Acalanes	San Francisco	1.20	0.74	0.23	37,939	0	0	0	40	115	32	45	36	11	11
Sales	Acalanes	San Jose	1.20	0.74	0.30	35,488	0	0	0	5	18	6	7	1	1	2
Sales	Alameda	Oakland	1.03	2.03	0.09	31,306	0	0	0	710	199	206	74	2	0	0
Sales	Alameda	San Francisco	1.03	2.03	0.20	37,939	0	0	0	54	92	44	55	50	0	0
Sales	Alameda	San Jose	1.03	2.03	0.26	35,488	0	0	0	28	47	19	28	3	0	0
Sales	Albany City	Oakland	1.20	1.17	0.09	31,306	0	0	0	184	17	3	12	0	0	0
Sales	Albany City	San Francisco	1.20	1.17	0.20	37,939	0	0	0	7	19	5	1	0	0	0
Sales	Albany City	San Jose	1.20	1.17	0.31	35,488	0	0	0	3	8	3	0	0	0	0
Sales	Alexander Valley	Santa Rosa	1.04	0.89	0.09	27,330	0	0	9	14	0	0	0	0	0	12
Sales	Alum Rock	Oakland	0.76	0.96	0.14	31,306	0	0	0	14	6	9	5	0	0	0
Sales	Alum Rock	San Francisco	0.76	0.96	0.21	37,939	0	0	0	2	16	14	24	34	13	4
Sales	Alum Rock	San Jose	0.76	0.96	0.11	35,488	0	0	0	90	307	290	309	270	22	1
Sales	Antioch	Oakland	0.93	0.81	0.16	31,306	0	0	0	836	290	445	434	6	8	0
Sales	Antioch	San Francisco	0.93	0.81	0.32	37,939	0	0	0	31	105	50	90	137	11	2
Sales	Benicia	Oakland	1.14	1.70	0.15	31,306	0	0	0	30	5	20	26	23	0	19
Sales	Benicia	San Francisco	1.14	1.70	0.28	37,939	0	0	0	6	21	5	9	23	8	12
Sales	Benicia	Vallejo	1.14	1.70	0.07	25,101	0	0	91	8	0	3	3	0	0	0
Sales	Berkeley	Oakland	0.98	1.35	0.08	31,306	0	0	0	613	91	85	31	6	0	0
Sales	Berkeley	San Francisco	0.98	1.35	0.20	37,939	0	0	0	41	73	18	22	3	0	0
Sales	Berkeley	San Jose	0.98	1.35	0.30	35,488	0	0	0	21	32	5	9	2	0	0

Table 4.14: Descriptive Statistics for 122 Housing Communities

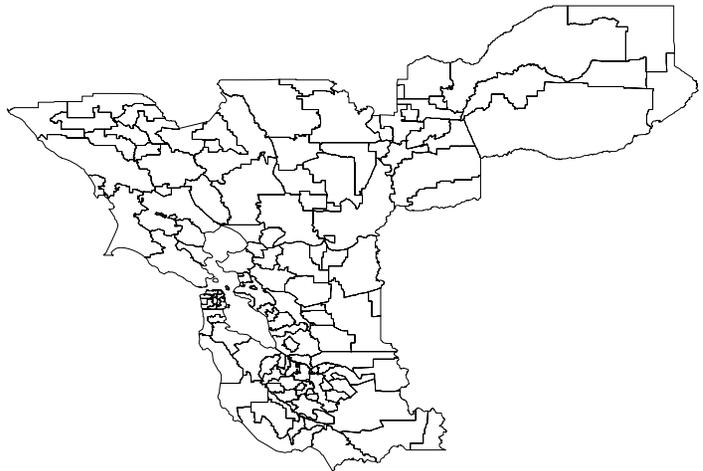
Observed Attribute	Source	Mean	St. Dev.	Min	Max
Community Size (population share)	Census	0.008	0.008	5.45E-05	0.047
Ozone (parts per million)	CA Air Resources Board	0.069	0.015	0.031	0.106
Academic Performance Index	CA Dept. of Education	706	93	528	941
Household Total Income (25 th quantile)	Census	46,047	14,133	22,291	104,137
Household Total Income (50 th quantile)	Census	74,779	23,065	41,977	174,591
Household Total Income (75 th quantile)	Census	115,016	35,368	62,759	239,195
Household Nonwage Income (25 th quantile)	Imputed	10,185	12,046	0	83,916
Household Nonwage Income (50 th quantile)	Imputed	29,565	17,339	5,109	96,792
Household Nonwage Income (75 th quantile)	Imputed	58,005	21,555	22,500	112,590
Housing Expenditures (25 th quantile)	<i>Dataquick</i>	27,825	12,565	9,156	88,082
Housing Expenditures (50 th quantile)	<i>Dataquick</i>	37,275	16,240	12,166	100,280
Housing Expenditures (75 th quantile)	<i>Dataquick</i>	48,345	21,127	16,407	123,624



* San Francisco, Sacramento, and Los Angeles are Consolidated Metropolitan Statistical Areas. All others are Metropolitan Statistical Areas. The unlabeled regions are individual counties that are not sufficiently urban to be considered part of a metropolitan area.

Figure 4.1: California's Metropolitan Statistical Areas*

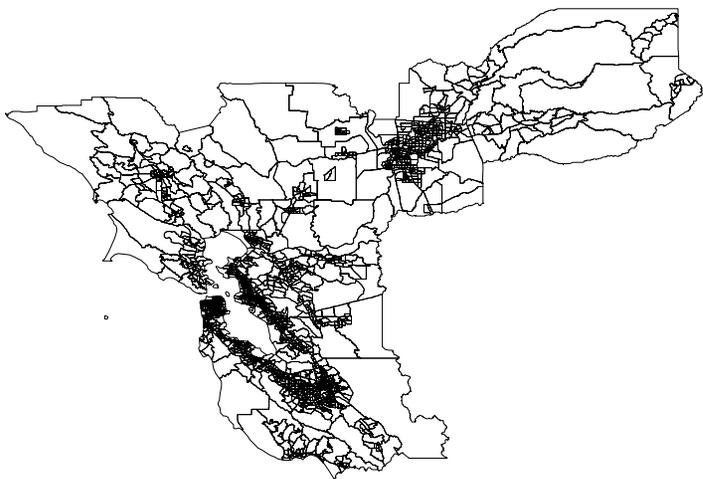
A: Housing Communities (School Districts)



B: Primary Metropolitan Statistical Areas



C: Census Tracts



D: School Districts & Air Quality Monitoring Stations

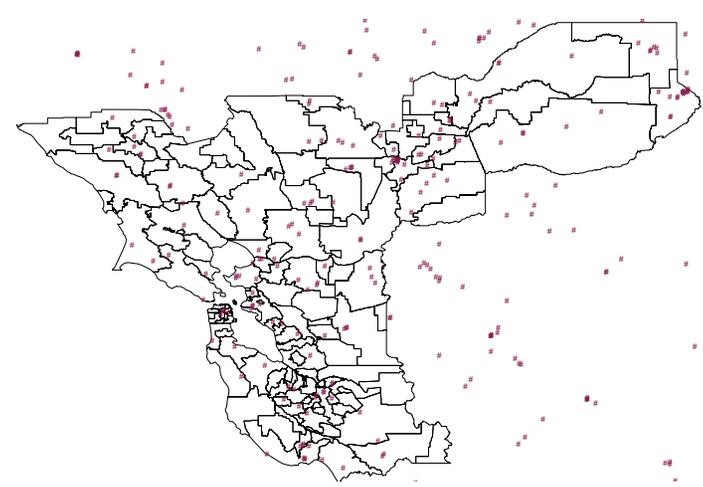


Figure 4.2: The Regional Landscape

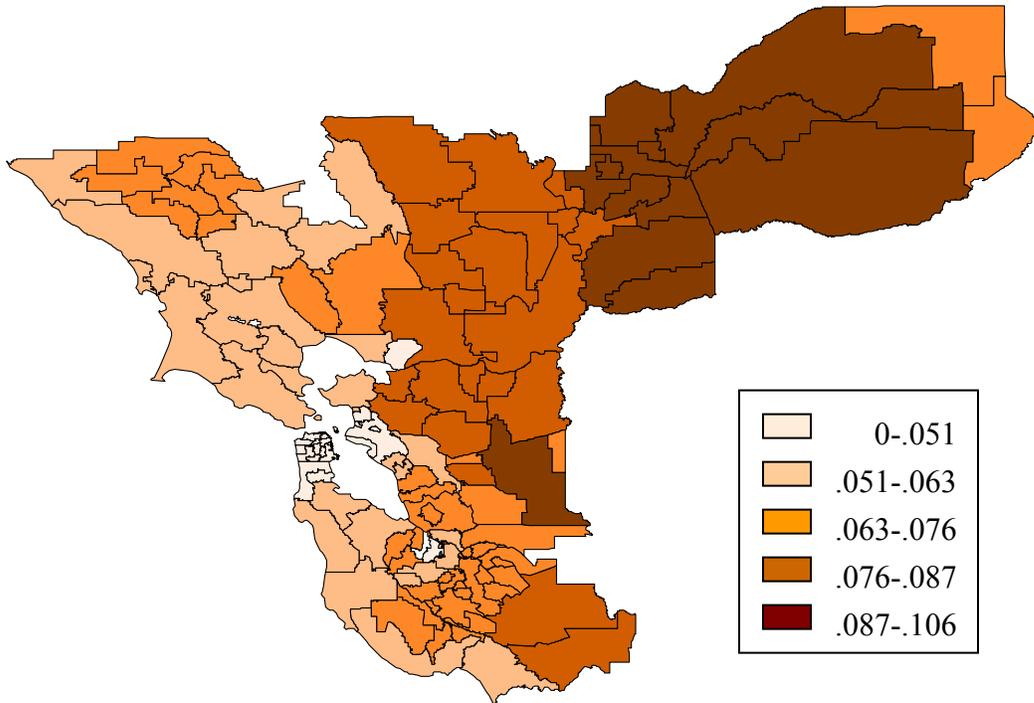


Figure 4.3A: Ozone by School District (parts per million)

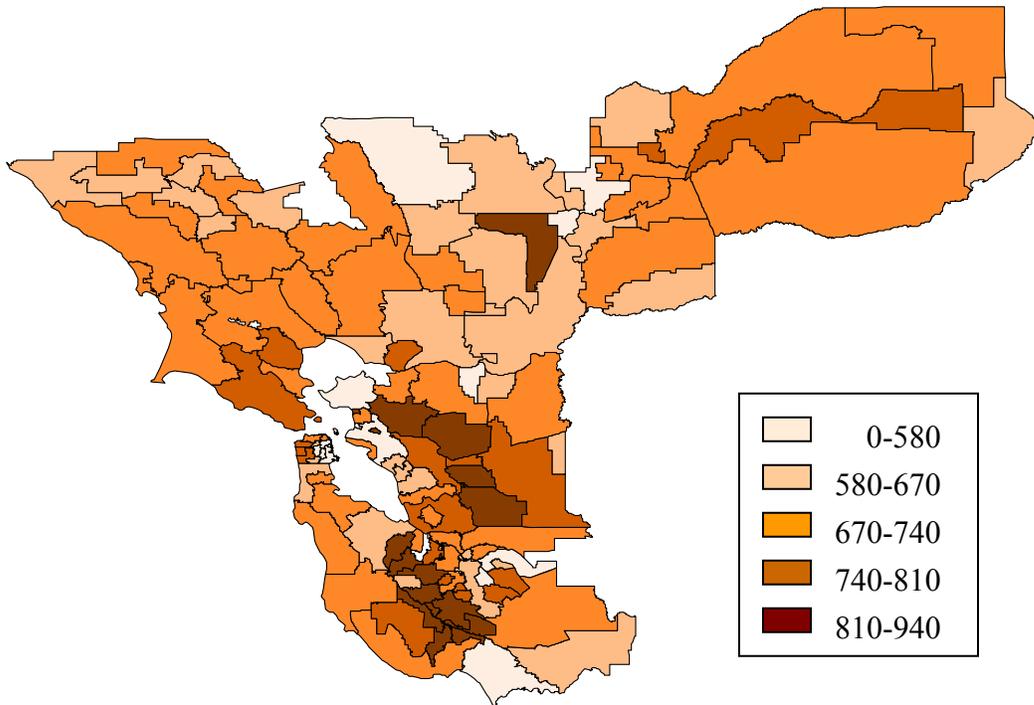


Figure 4.3B: Academic Performance Index by School District

Chapter 5: Estimation Results

I. Introduction

This chapter estimates the relationship between households' location choices in the San Francisco-Sacramento region of California and community-specific provision of local public goods, recognizing that a household's wage income will depend on its location. The estimation process uses data on each household's income, the occupation of its primary earner, and the attributes of its location choices to recover values for the parameters of the indirect utility function that is hypothesized to explain its choice of housing community and work destination. These parameter values can be used to describe the household's demand for two local public goods: air quality and school quality. While both are presumed to influence location choices, air quality is the focus of the policy application in chapter 6. Therefore, the estimation results are analyzed in terms of their implications for the willingness-to-pay for improved air quality.

Recall from chapter 3 that the empirical model generalizes the "Epple-Sieg" sorting framework to: (1) recognize that wage income varies with location, (2) recover horizontally differentiated preferences for local public goods, and (3) relax the standard practice of making specific assumptions about the distribution of the preference parameters used to characterize sources for unobserved heterogeneity. In addition to using the new model to explain observed behavior, this chapter aims to evaluate the economic implications of these generalizations. This second aspect of the research is addressed by comparing results from the new "dual-market" estimator to the results from two special cases—the Epple-Sieg model, and an intermediate version of the estimator that admits horizontal differentiation but treats wage income as exogenous.

Section II begins with an overview of the estimation process and contrasts the three versions of the model. Each is based on a similar expression for indirect utility, and thus they share some common parameters. Section III reports the results from estimating an index of housing prices and the demand for the private good component of housing, both of which are common to all three versions of the model. These results are then treated as known constants during the second stage of the estimation which recovers a distribution of heterogeneous parameters that reflect preferences for local public goods. Section IV

describes results for the Epple-Sieg specification, and section V describes results for the other two versions of the model in a way that parallels the development of the econometric model in chapter 3. Finally, section VI compares the economic implications of the results across the three models.

II. Overview of the Estimation

Together, the three versions of the model provide the means to evaluate the economic implications of introducing horizontally differentiated preferences and job opportunities into the Epple-Sieg sorting framework. The most general version is the dual-market approach which depicts households as having horizontally differentiated preferences for public goods and making a joint job-house choice. The term “dual-market” refers to there being two *types* of markets, community-specific markets for housing and labor markets. In this case, the choice set consists of the 268 (housing community, labor market) combinations and a household’s income will vary across locations depending on the occupation of its primary earner and the required commute time. This framework nests the other two versions of the model as special cases. In the intermediate “single-market horizontal” case, income is treated as exogenous to location choice so that households only choose among the 122 housing communities. Finally, the “single-market vertical” case treats income as exogenous, preferences as vertically differentiated, and restricts the joint distribution of income and preferences to be lognormal. This formulation corresponds to the Epple-Sieg model.

The three versions of the model can be related in terms of the indirect utility function from equation (3.1), which is repeated here for convenience:

$$(5.1) \quad V_{i,s,z} = \left\{ \alpha_i (\bar{g}_{i,z})^\rho + \left[\exp\left(\frac{y_{i,s,z}^{1-\nu} - 1}{1-\nu}\right) \exp\left(-\frac{\beta p_z^{\eta+1} - 1}{1+\eta}\right) \right]^\rho \right\}^{\frac{1}{\rho}},$$

where $\bar{g}_{i,z} = \gamma_{i,1} g_{1,z} + \dots + \gamma_{i,N-1} g_{N-1,z} + \gamma_{i,N} \bar{\xi}_z$, and

$$y_{i,s,z} = \hat{y}_i + \theta_{i,1} w_{s,z} (1 - \theta_{i,2} t_{s,z}).$$

This expression for utility forms the basis for the dual-market version of the estimator. In the single-market horizontal case, the job skill parameter (θ_1) is restricted to equal 1 and the opportunity cost of time parameter (θ_2) is restricted to equal 0. These same restrictions are imposed in the single-market vertical case, which also assumes that $f(\alpha, y)$ is lognormally distributed, and requires the weights in the public good index to be constant across households so that the i subscript is dropped from the γ 's.

Figure 5.1 summarizes the maintained assumptions in each version of the model and outlines the estimation strategy. Recall that the estimator developed in chapter 3 can be decomposed into two stages. The first stage recovers the price of a homogeneous unit of housing in each community (p_1, \dots, p_{122}), and the homogeneous housing demand parameters (β, η, ν). These results are treated as known constants during the second stage of the estimation, which recovers a composite unobserved public good for each community ($\bar{\xi}_1, \dots, \bar{\xi}_{122}$), the homogeneous elasticity of substitution parameter (ρ), and a partition of preference space for the heterogeneous parameters $A(\alpha, \gamma, \theta)$.

Since the differences between the three versions of the model do not affect p_1, \dots, p_J and β, η, ν , the first stage of the estimation was only performed once. Likewise, the second stage of the estimation was performed simultaneously for the two horizontally differentiated versions of the model. More precisely, the same estimates for ρ and $\bar{\xi}_1, \dots, \bar{\xi}_{122}$ were used to recover an approximation to the partition preference space for the single and dual-market models. The only difference is that two additional dimensions of preference space were partitioned in the dual-market case (θ_1, θ_2). This isolates the way that including job opportunities in the model affects the resulting partition of preference space. Finally, in the single-market vertical case, the second-stage parameters were estimated using the GMM approach developed by Sieg et al. (2004). Comparing the results to those from the two horizontally differentiated models demonstrates the implications of relaxing vertical differentiation and the lognormal assumption on the joint distribution of income and preferences.

III. Estimation: First-Stage Results

In the first stage of the estimation, the data on individual real estate transactions were used along with income distributions for each community to estimate an index of housing prices and the homogeneous housing demand parameters. Part (A) of this section describes how the index of housing prices was estimated. Then part (B) explains how the resulting index was used together with the income data to recover the housing demand parameters.

(A) *The Price of Housing* (p_1, \dots, p_J)

Recall from chapter 3 that community-specific fixed effects estimated from a hedonic price function can be used to construct a theoretically consistent index for the price of a homogeneous unit of housing in each community. More specifically, given a separability restriction on preferences, expenditures on housing can be factored into the product of a price index and a quantity index⁵⁵. The first equality in (5.2) displays this result.

$$(5.2) \quad \ln e_{j,n} = \ln p_j + \ln \bar{h}(h_n) = \ln p_j + [\sigma \cdot \ln h_n + \delta \cdot \ln h_n' h_n + \psi \cdot \ln h_n h_n'] + \mu_{j,n}.$$

Log expenditures on house n in community j equal the log price of housing in that community plus the log “quantity” of housing consumed, which depends on h_n , a vector of structural characteristics describing house n . Using this relationship recover the price index requires selecting a parameterization for the quantity index. The second equality in (5.2) shows how this was done. The term in brackets expresses the quantity of housing using a log-linear index of the housing characteristics, the characteristics squared, and interactions between the characteristics⁵⁶. This particular specification follows Sieg et al. (2002, 2004).

⁵⁵ The restriction is that structural characteristics of housing enter the direct utility function through a sub-function that is HOD 1 and separable from the effect of public goods and the numeraire.

⁵⁶ Linear and semi-log specifications were also estimated in order to test the sensitivity to functional form assumptions. This had very little impact on the resulting price index, mainly affecting the index numbers for the most expensive communities. Correlation coefficients between indices were 0.99 for the semilog and the

Equation (5.2) was estimated by OLS using 540,642 observations on the sale prices of homes and the number of bedrooms, number of bathrooms, lot sizes, building sizes, age of each house, and a dummy variable for condominiums (see table 4.3 for summary statistics). The resulting coefficients and their standard errors are shown in table 5.1. Most of the coefficients are statistically significant. The R^2 of 0.81 indicates that the structural characteristics and community-specific fixed effects explain most of the variation in housing prices. Negative signs for the coefficients on characteristics that would normally be expected to increase the sale price of a home (e.g. lot size) probably reflect the measure used, collinearity, or both.

The community-specific fixed effects recovered from the regression indicate that housing in the most expensive community costs 6.5 times as much as in the cheapest community. After normalizing by the lowest price, the index ranges from 1.00 in Sacramento's Grant Union high school district to 6.51 in San Francisco's second supervisorial district⁵⁷. This range is comparable to the results from previous sorting applications. For example, using the same expression for the hedonic expenditure function, Sieg et al. (2004) find that prices range from 1.0 to 7.0 in the Los Angeles CMSA.

Figure 5.2 illustrates the implied distribution of prices. The mean price is 2.55 and the standard deviation, 1.08. Overall, the distribution is consistent with the conventional wisdom that the San Francisco Bay Area is an expensive place to live. The price of housing for the median community in the San Francisco CMSA is 2.62. In contrast, the most expensive community in the Sacramento CMSA (Davis Joint Unified) has a price of only 2.08. To help illustrate the spatial variation in prices, figure 5.2 uses different symbols to identify the PMSA in which each community is physically located. Spatial correlation is most evident in the tails of the distribution. The 11 cheapest communities are all located in the Sacramento PMSA, while 24 of the 25 most expensive communities are in the San Francisco and San Jose PMSAs.

Despite the spatial concentration of communities with extreme values for the price

current (double-log) specification, and 0.96 for the linear and double-log.

⁵⁷ San Francisco's 2nd supervisorial district comprises the area just southeast of the Golden Gate Bridge, bordering the Pacific Ocean. It includes the city's affluent Marina district.

index, there is considerable variation within most of the PMSAs. For example, prices in the Oakland PMSA range from 1.42 in Pittsburg Unified to 4.23 in Piedmont City Unified. This is illustrated in table 5.2, which shows the median, minimum, and maximum prices for each PMSA. The price of housing varies by more than 100% between the most expensive and least expensive communities in Oakland, San Francisco, San Jose, and Vallejo⁵⁸.

Furthermore, price ranges within many of the PMSAs overlap. For example, the price index has a value of 1.64 in the Woodland Joint Unified school district, which is located in the Yolo PMSA. A household living there could move to communities in Sacramento, Vallejo, Santa Rosa, or Oakland and pay a similar price for housing. The price index was used together with data on the distribution of income and housing expenditures in each community to estimate the demand for the private good component of housing.

(B) *Housing Demand Parameters* (β, η, ν)

The homogeneous housing demand parameters were estimated from the housing expenditure function in equation (5.3).

$$(5.3) \quad \ln(e_j^{quant}) = \ln(\beta) + (\eta + 1)\ln(p_j) + \nu \ln(y_j^{quant}) + \varepsilon_j.$$

Specifically, quantiles from the distribution of annualized housing expenditures in each community were regressed on the price of housing and quantiles from the income distribution. The 25th, 50th, and 75th quantiles were used. Recall from chapter 3 that there is reason to expect both prices and income to be endogenous in the regression. By construction, the price index will be correlated with the unexplained component of housing expenditures. Likewise, income may be correlated with the error term if it is endogenous to location choice. Therefore the expenditure function was estimated using 2SLS in addition to OLS. The 2SLS regression used the observed public goods as instruments for the price of housing and the 25th, 50th, and 75th quantiles from the distribution of non-wage income as

⁵⁸ Prices range from 2.74 to 6.51 between the city of San Francisco's 11 supervisorial districts. This supports the decision to subdivide the city into supervisorial districts discussed in chapter 4.

instruments for total income. Table 5.3 reports the regression results⁵⁹.

The first three rows of table 5.3 report the results from OLS and 2SLS estimation of the expenditure function. Including instruments in the regression produces a modest increase in the income elasticity and a modest decrease in the price elasticity relative to OLS. As the elasticities increase in absolute magnitude the demand intercept decreases. The estimates for the price elasticity are similar to the results from previous sorting applications. For example, the 2SLS estimate ($\hat{\eta}_{2SLS} = -0.35$) falls near the middle of the range reported in the existing literature (-0.01 to -0.70).⁶⁰ In contrast, the estimates for the income elasticity are relatively low. The largest estimate ($\hat{\nu}_{2SLS} = 0.57$) falls below the range of results from previous studies (0.73 to 0.94). The simplest explanation is that the demand for housing is less sensitive to income in San Francisco and Sacramento than in the metropolitan areas that have been the subject of previous sorting applications (e.g. Los Angeles, Boston, Portland). However, earlier studies that used national data to estimate the demand for housing produced nearly the same range of estimates. In his summary of this work, Polinsky (1977) reports that consistent estimates of the income elasticity range from 0.75 to 0.90. A second explanation for why the current estimates fall outside this range stems from differences in the way income is measured.

While the demand for housing is presumed to depend on permanent income, the expenditure function in (5.3) was estimated using data on current income. In contrast, the earlier studies summarized by Polinsky (1977) used an approximation to permanent income⁶¹. Likewise, some of the larger estimates for ν in the sorting literature come from specifications in Sieg et al. (2004) where permanent income is treated as a latent variable and

⁵⁹ Table 4.14 reports summary statistics for the data used in the estimation.

⁶⁰ This includes all sorting applications that have estimated (5.2) directly or included it as a moment condition in GMM estimation: Epple and Sieg (1999), Walsh (2003), Wu and Cho (2003), Sieg et al. (2004), and Epple, Peress and Sieg (2005). Polinsky (1977) reports a lower range of estimates (-0.87 to -0.67) in his summary of consistent micro models. Unlike the sorting literature, these earlier studies typically defined the price of housing using a (Bureau of Labor Statistics) index that does not control for variation in the structural characteristics of homes.

⁶¹ The studies summarized in Polinsky (1977) estimated the demand for housing using data on households who had obtained mortgages from the Federal Housing Administration (FHA). They defined a household's permanent income using FHA's estimate of the household's income during the first third of their mortgage.

they estimate the parameters of its (lognormal) distribution.

In general, when data on current income are used to estimate the demand for housing, the resulting income elasticity will be biased downward if current income tends to fall below permanent income in expensive communities. This type of systematic difference can be caused by appreciation in real estate values. Ortalo-Magné and Rady (2006) provide the following example:

“There are neighborhoods in central London where old taxicab drivers live next to young investment bankers. The taxicab drivers did not move to the neighborhood at the same time as the investment bankers. Rather, they bought their home several years before the investment bankers moved in. As rents went up in their neighborhood, so did their wealth.”

This story may also apply to communities in the San Francisco Bay Area where housing prices have increased dramatically over the past 20 to 30 years. To a lesser extent, this may be true for the study region as a whole. For example, between 1984 and 2004 the (Bureau of Labor Statistics) index of housing prices increased by 111% in the San Francisco CMSA compared to an 84% average increase for all U.S. cities.

One way to lessen the extent to which differences between current and permanent income affect estimates for the income elasticity is to exclude retired households from the distribution of current income. It seems reasonable to expect that, for retired households, permanent income will generally exceed current income. This effect will be magnified in communities where the price of housing has appreciated substantially, placing a downward bias on estimates for the income elasticity. Under the maintained assumption that all households share the same demand for housing, excluding retired households should not affect the consistency of parameter estimates. That is, since all households are assumed to share the same values for β , η , and ν , consistent estimates of these parameters can be obtained by estimating the expenditure function using any subset of the data.

The last three rows of table 5.3 report the results from OLS and 2SLS estimation of the expenditure function using the distribution of current income for working households. As expected, excluding retired households increases the income elasticity. While the 2SLS

estimate ($\hat{v}_{2SLS} = 0.66$) is still slightly lower than the range of previous results, their 95% confidence intervals overlap. Aside from the larger income elasticities, the pattern of results is similar to when retired households are included in the estimation

The distinction between permanent and current income could be addressed more directly with access to the restricted census micro data that were discussed in the conclusion to the previous chapter. Although the micro data do not include a direct measure of wealth or permanent income, they include detailed information on each household's sources of income (e.g. wages, dividends, rental income). This could be used together with a household's demographic characteristics to construct an estimate for its permanent income. This task is left for future research. For the remainder of this chapter the demand for housing is defined using the 2SLS estimates reported in bottom row of the table 5.3: $\beta = 11.97, \eta = -.38, \nu = .66$.

IV. Estimation: Second-Stage Results for the Vertical (Epple-Sieg) Model

If wage income is exogenous, households have vertically differentiated preferences for public goods, and the shape of the joint distribution of income and preferences is known, then the second-stage of the estimation is relatively straightforward. Under those conditions, all the parameters of the model can be estimated simultaneously using the GMM approach developed by Sieg et al. (2004). Part (A) of this section provides a brief overview of the GMM estimation strategy, before discussing results in part (B).

(A) Overview of the GMM Approach to Estimation

Treating the first-stage estimates for the housing demand parameters as known constants, the GMM estimator can be used to recover the CES parameter (ρ), the parameters that characterize the joint lognormal distribution of income and preferences ($\mu^\alpha, \mu^y, \sigma^\alpha, \sigma^y, \lambda$), the overall level of public goods provision in the cheapest community (\bar{g}_1), and the constant weight on air quality in the public goods index (γ_{air}). Let ψ represent a vector of these

parameters. Equation (5.4) defines the GMM objective function, where z is a set of instruments, m represents the moment conditions, and V is the covariance matrix of moments.

$$(5.4) \quad \psi = \arg \min_{\psi \in \Psi} \left\{ \frac{1}{J} \sum_{j=1}^J z_j m_j(\psi) \right\}' V^{-1} \left\{ \frac{1}{J} \sum_{j=1}^J z_j m_j(\psi) \right\}.$$

Sieg et al. demonstrate that the seven moment conditions in (5.5) can be used to identify all the parameters in ψ .

$$(5.5) \quad m_j(\psi) = \left\{ \begin{array}{l} \bar{g}_j - 1 \cdot school - \gamma_{air} \cdot air \\ y_j^{25} - \tilde{y}_j^{25} \\ y_j^{50} - \tilde{y}_j^{50} \\ y_j^{75} - \tilde{y}_j^{75} \\ \ln e_j^{25} - \ln \beta - (\eta + 1) \ln p_j - \nu \ln \tilde{y}_j^{25} \\ \ln e_j^{50} - \ln \beta - (\eta + 1) \ln p_j - \nu \ln \tilde{y}_j^{50} \\ \ln e_j^{75} - \ln \beta - (\eta + 1) \ln p_j - \nu \ln \tilde{y}_j^{75} \end{array} \right\}.$$

The first condition is based on the level of public goods provision. Given a value for overall provision of public goods in the cheapest community, \bar{g}_1 , the sorting behavior implied by vertical differentiation allows $\bar{g}_2, \dots, \bar{g}_{122}$ to be defined recursively. The predictions for $\bar{g}_1, \dots, \bar{g}_{122}$ are then used to identify the (constant) weight on air quality in the public goods index while the weight on school quality is normalized to equal 1. The residual to the moment condition defines the composite unobserved public good in each community ($\bar{\xi}_1, \dots, \bar{\xi}_{122}$).

The other six moment conditions are based on the model's prediction for the distribution of income in each community, $F(\tilde{y})$. Under the maintained assumptions on preferences, the information in ψ can be used to simulate community-specific income

distributions. Three of the moment conditions match the 25th, 50th, and 75th quantiles from the simulated distributions of income in each community with the corresponding quantile from the observed income distributions. The other three moment conditions use the simulated income distributions to match predicted and observed quantiles from the distribution of housing expenditures in each community.

Instruments are required to address endogeneity in the moment condition based on provision of public goods. The problem is that unobserved public goods will be correlated with overall public goods provision in each community. This can be addressed using instruments created from monotonic functions of each community's rank in the price index. These instruments will provide the necessary identification if unobserved public goods are of second order importance; i.e. if they affect households' location choices without affecting the price rank of a community (Sieg et al. [2004]). In general, there is also reason to believe that prices and income will be endogenous in the housing expenditure function, as discussed in the previous section. However, this is not a problem for the current application because the first-stage estimates for β, η , and ν are treated as known constants during the GMM estimation. This two-step approach is the main difference between the current application and the estimation in Sieg et al. (2004) where β, η , and ν are included in ψ .

(B) Results from GMM Estimation

Epple and Sieg (1999) prove that all of the parameters of the (vertically differentiated) model are locally identified. However, they do not appear to be globally identified in the current application. In other words, parameter estimates appeared to be very sensitive to the choice of starting value used for the GMM estimation. Therefore, a global optimization algorithm was used to search for a good starting value during the first stage of the GMM estimation. Specifically, the genetic algorithm described in appendix B was used to search over the bounded space defined in columns 1 and 2 of table 5.4. The bounds on each parameter were chosen using one of two strategies. For the parameters which had fairly consistent point estimates during the initial attempts at estimation, bounds were chosen to include the range of observed results. For example, all the initial results for the mean of the log income

distribution fell between 10.9 and 11.1. Likewise, all the point estimates for the CES parameter fell between -0.01 and -0.10. The genetic search used slightly wider bounds for each of these parameters. Alternatively, for parameters that appeared to be sensitive to the choice of starting value, bounds were chosen to include the range of results reported for the comparable specifications in Sieg et al. (2004). For example, while Sieg et al. report values for μ^α between 0.41 and 0.83, the genetic search used a range from 0.0 to 1.0. Finally, it is important to emphasize that the bounds shown in the table were only used to find a starting value for the first-stage of the GMM estimation. They did not constrain the estimation in any other way.

The genetic algorithm produced a starting value near a local minimum⁶². This can be seen by comparing the starting value for each parameter (table 5.4, column 3) with its corresponding point estimate from the subsequent GMM estimation (column 4). This is not surprising since the genetic algorithm is designed to search for the global minimum of a function.

Most of the parameters in column 4 of table 5.4 are precisely estimated and similar in magnitude to the results in Sieg et al. The negative correlation between income and preferences for public goods ($\lambda < 0$) reflects the fact that there is considerable overlap in the community-specific income distributions. Alternatively, if λ were positive, the model would predict almost no overlap in the range of income within different communities. The slightly negative value for ρ indicates the elasticity of substitution between public and private goods is slightly less than one, which implies the marginal willingness-to-pay for public goods is increasing in income. This is consistent with the single-crossing restriction on preferences, providing a consistency check on the theoretical model.

Recall from chapter 2 that in a vertically differentiated model households can be ordered along an interval according to their preferences for public goods. For example, figure 5.3 displays the implied ordering for households with income equal to \$50,000. A

⁶² The starting value was the “best” point the genetic algorithm found in 300,000 evaluations of the first-stage GMM objective function, with the weights matrix defined as an identity matrix. This took approximately 20 hours on a standard desktop computer. The GMM objective function used a simulated “population” size of 100,000 households.

household with $\alpha < 0.69$ will maximize its utility by purchasing 15,115 units of housing in Grant Joint Union high (the lowest priced community), while a household with the same income and $\alpha > 21.71$ will purchase 7,411 units of housing in San Francisco's second supervisorial district (the highest priced community)⁶³. The positive coefficient on air quality in the public goods index indicates that, all else held constant, households with higher values for α will be willing to pay more for a marginal improvement in air quality.

Point estimates for the parameters that reflect preferences for public goods have the expected signs, but relatively large standard errors. This is true for μ^α , \bar{g}_1 , and especially γ_{air} . Since all three are identified by the moment condition based on public goods provision, the large standard errors could signal a problem with the rank instruments⁶⁴. For example, some of the unobserved public goods may be as important, or more important, than air quality and school quality in determining where households choose to live. If so, $\bar{\xi}$ could have a first-order effect on the price ranking of communities, violating the orthogonality requirement. Such violations seem likely given that the average community's ranking by $\bar{\xi}$ differs from its price ranking by only 8 places.

Problems with the rank instruments could be addressed in one of four ways. First, additional public goods could be added to the index. For example, data could be collected on crime, climate, access to open space, distance to major cities, and other amenities that seem likely to be reflected by $\bar{\xi}$. This would diminish the contribution of $\bar{\xi}$ to a community's price rank at the cost of increasing the number of index weights that need to be identified during the GMM estimation. Alternatively, if households are free to choose approximately continuous quantities of the additional public goods, then their contribution to the housing price index can be removed during the first-stage of the estimation. For example, Sieg et al. (2004) include variables for *distance to the coast* and *commuting distance* in the first-stage hedonic regression. A third way to address problems with the rank instruments would be to look for different instruments that satisfy the orthogonality requirement. Finally, a fourth

⁶³ The "units" of housing purchased were calculated from the housing demand function: $\bar{h} = \beta p_j^\eta y_i^\nu$.

⁶⁴ The appendix to Epple and Sieg (1999) explains how each parameter in the model is identified.

possibility would be to adapt the strategy used to recover $\bar{\xi}$ in the horizontal version of the estimator. The idea would be to use the kernel approach described in chapter 3 to recover a set of values for $\bar{\xi}$ from the joint distribution of prices, air quality, and school quality. The resulting estimates could be treated as an additional (observed) public good during the GMM estimation. In this case, the residual to the moment condition based on public goods provision could be interpreted as measurement error. An advantage of the kernel approach is that it does not place any restrictions on the relative contributions of observed and unobserved public goods to the price index. The next section reports results from implementing this strategy for the horizontal model and speculates on the amenities that seem likely to be reflected by $\bar{\xi}$.

Relaxing vertical differentiation and a priori assumptions about the distribution of heterogeneous preferences would make it difficult, if not impossible, to recover the second-stage parameters using GMM-type estimation. When households are allowed to differ in their relative preferences for the various public goods in the index, they will also differ in how they perceive composite provision of public goods in each community. This makes it impossible to predict unique values for $\bar{g}_1, \dots, \bar{g}_{122}$ as required by the first moment condition in (5.4). Moreover, without a priori distributional assumptions it is no longer possible to simulate the community-specific income distributions that formed the basis for the other six moment conditions in (5.4). As a result, the horizontally differentiated model described in the next section adopts a different estimation strategy.

V. Estimation: Second-Stage Results for Horizontal Models

In addition to allowing horizontally differentiated preferences and relaxing the need for a priori distributional assumptions, the econometric model outlined in chapter 3 recognizes that working households make a joint job-house choice. Nonetheless, it restricts the data generating process and can be computationally intensive to implement. This section describes how these issues were addressed and reports the results from using the model to

simultaneously estimate all the second-stage parameters. Part (A) provides an overview of the estimation strategy and discusses implementation issues. Then parts (B), (C), and (D) describe the resulting estimates for ρ , $\bar{\xi}$, and the partitions of preference space that were recovered for single and dual-market versions of the model.

(A) *Overview of the Second Stage Estimation Strategy*

Recall that the estimator uses an iterative process to simultaneously recover all the second-stage parameters. The iterative structure (depicted in figure 3.5) is based on solving for a point estimate of ρ . On the first iteration, a starting value (ρ^0) is used to solve for a vector of unobserved public goods ($\bar{\xi}_1^0, \dots, \bar{\xi}_{122}^0$) which are then used together with ρ^0 to partition preference space. The resulting partition is used to evaluate an objective function that equals zero at the true value of ρ (i.e. equation (3.14)). Then, the value of the objective function is used to choose a new value for the CES parameter (ρ^1) to be used during the second iteration. This process terminates when additional changes in ρ do not lead to further improvements in the objective function.

Implementing the estimator requires addressing two issues. First, identifying the true value of ρ requires preferences for public goods to be independent of income. This restriction is not likely to be satisfied for all households. For example, children in private schools tend to come from higher-income families. The average income of households with children enrolled in private schools in the San Francisco-Sacramento area is 42% higher than for those enrolled in public schools (*Census School District Special Tabulation*, 2000). These choices would seem to imply we should expect a negative correlation between income and strength of preferences for local public school quality.

The second issue is the computational burden associated with recovering preferences. Chapter 3 developed a Gibbs sampling algorithm to draw points uniformly from the region of preference space that rationalizes each location choice. Figure 5.4 illustrates how the resulting sample of points provides an approximation to the partition of preference space. Panel A displays the partition from the 4-community example discussed in the previous

chapters, and panel B displays an approximation to that partition based on a sample of points drawn uniformly from each region. Three features of the current application make it computationally intensive to produce an approximation like the one given in panel B. First, with n heterogeneous preference parameters, taking a single uniform draw requires solving $2n$ nonlinear rootfinding problems⁶⁵. Second, it may require a large number of draws to produce an accurate representation of each region. Third, the partition has a large number of regions. More precisely, because the indifference loci that delineate the regions of preference space are a function of both wages and non-wage income, the partitioning process must be repeated at every level of non-wage income for workers in each occupation. For example, in the dual-market version of the model there are 22 occupations and households with primary earners in each occupation are classified according to 10 different levels of non-wage income. With 268 job-house locations, the partition contains $22 \times 10 \times 268$ different regions, each of which has 6 dimensions—one for each heterogeneous preference parameter. Drawing 1000 points from every region requires solving over half a billion nonlinear rootfinding problems. Moreover, the partitioning process must be repeated on every iteration of the objective function used to solve for ρ .

The maintained assumption that everyone shares the same values for ρ and $\bar{\xi}$ forms the basis for a strategy to simultaneously satisfy the independence restriction and reduce the computational burden. Suppose the independence restriction holds for a known subset of households. Using only those households, the (iterative) estimation can be performed to obtain consistent estimates for ρ , $\bar{\xi}$, and an approximation to the partition that rationalizes their location choices. Then, treating the estimates for ρ and $\bar{\xi}$ as known constants, preference space only needs to be partitioned once for the remaining households. This strategy was used to recover ρ and $\bar{\xi}$ from data on retired households.

Retired households were a strategic choice for two reasons. First, they are less likely to violate the independence assumption required to identify ρ . Since they are less likely than working households to have school-age children, there is less reason to expect their

⁶⁵ The rootfinding problem are defined in chapter 3.V.

preferences for public school quality to be correlated with their income. There is also no obvious reason to expect correlation between their income and preferences for other public goods. For example, poor air quality should affect retirees' health regardless of income.

The second strategic advantage of using retired households is that they bridge the single and dual-market versions of the model. Generalizing the urban landscape to include labor markets does not affect the choice set faced by retirees; i.e. their income is always fixed. Since retirees choose from the same 122 communities in both versions of the model, both models should return the same information about their preferences. This requires both models to produce the same estimates for ρ and $\bar{\xi}$, which is guaranteed if they are estimated from data on retired households. Furthermore, retired households impose a binding constraint on the values assigned to the unobserved public good. In other words, any $\bar{\xi}$ -vector that allows the model to recover preference sets that rationalize retired households' location choices will also allow it to explain the choices made by working households, whereas the opposite condition is not true⁶⁶.

Finally, a caveat to using retired households is that systematic differences between their current and permanent income could bias the estimate for ρ . Recall that the strategy for recovering the housing demand parameters in the first stage of the estimation presumed that location choices depend on permanent income and that, for retired households, permanent income exceeds current income. Under those assumptions, using data on current income to partition preference space will place an upward bias on α .⁶⁷ This may or may not affect the resulting estimate for ρ . If the bias on α is uncorrelated with the level of income, ρ should be unaffected since it is identified by minimizing the difference between the distributions of α for households with different income levels. Conversely, if the bias on α is correlated with income, ρ should be affected. One potential source of such correlation would be if housing prices had appreciated substantially in the most expensive communities. This could cause the difference between current and permanent income to be an increasing

⁶⁶ This is proven in lemma 2 of appendix A.

⁶⁷ All else held constant, the estimator assigns lower values for α to higher-income households.

function of income—the same problem that motivated the decision to exclude data on retirees while estimating the demand for housing. Since households in the most expensive communities will also tend to be assigned the highest values for α , this problem is most likely to occur the right tail of the distribution.

(B) *CES Parameter (ρ)*

Retired households were classified according to 10 income “bins” reported in the Census data, and each household was assigned a level of income equal to the midpoint of its bin⁶⁸. Then, the objective function used to estimate ρ was defined as the sum of the difference in the marginal distributions of $F(\alpha, \gamma_{air}, \gamma_{school}, \gamma_{\xi})$ for all pairwise combinations of income. To calculate the objective function, the marginal distributions were evaluated at their deciles from 0.1 to 0.9. Excluding the tails of the distributions reduces the sensitivity to extreme values and, in the case of α , to bias caused by appreciation in housing prices in the most expensive communities.

The objective function was minimized using a grid search over $[-.6, .1]$, which includes the range of estimates from previous studies. Figure 5.5 graphs the objective function over that range. While there appear to be multiple local minima, most are located between -0.15 and -0.10 . The function was minimized at $\rho = -0.14$. As in the vertically differentiated model, this provides a consistency check on the theoretical model since $\rho < 0$ guarantees that marginal willingness-to-pay for public goods is increasing in income. While the estimate for ρ is more than 6 times as large as the result from the vertical model (0.022), they imply similar values for the elasticity of substitution between public and private goods. Here, the elasticity is 0.88 compared to 0.98 in the vertical case.

(C) *Composite Unobserved Public Good ($\bar{\xi}_1, \dots, \bar{\xi}_J$)*

For each of the 122 communities, a value for $\bar{\xi}$ was estimated from the joint distribution of

⁶⁸ Measured in thousands, the midpoints are: [5 12.5 22.5 35 45 55 67.5 87.5 112.5 175].

housing prices, air quality, and school quality. Recall from chapter 3 that the estimator places two restrictions on the data generating process. Housing prices must be increasing in $\bar{\xi}$, and $\bar{\xi}$ must be generated independently of the observed public goods. Assuming these restrictions hold, the first equality in (5.6) illustrates the theoretical relationship that motivates the estimator.

$$(5.6) \quad F_{\bar{\xi}}(\bar{\xi}_j) = F_{p|g=g_j}(p_j) = \bar{\xi}_j.$$

The quantile of $\bar{\xi}_j$ in its distribution will equal the quantile of p_j in its distribution, conditional on air and school quality. Since $\bar{\xi}$ is only defined up to a monotonic transformation, it can be normalized to equal its distribution function, as illustrated by the second equality in (5.6). Thus, the normalized distribution of unobserved public goods can be recovered by estimating $F_{p|g=g_j}(p_j)$ nonparametrically from the index of housing prices and the data on air and school quality.

In addition to accounting for the presence of unobserved public goods, $\bar{\xi}$ is required to “pick up the slack” in explaining observed behavior. More precisely, the estimate for $\bar{\xi}_j$ must enable the analyst to find a set of values for the heterogeneous preference parameters that make community j the utility-maximizing location. This requirement can be satisfied by constraining the bandwidth on a kernel estimation of $F_{p|g=g_j}(p_j)$ to induce a sufficient amount of “smoothness” in the relationship between prices and $\bar{\xi}$.

By guaranteeing the model can rationalize every observed location choice, the approach to estimating $\bar{\xi}$ simultaneously accounts for the presence of public goods other than air and school quality and addresses restrictions imposed by the shape of the indirect utility function (i.e. misspecification). These two roles can be disentangled under the assumption that the true form of the indirect utility function is *supermodular* in public goods

and $1/p$.⁶⁹ If utility is supermodular, any community that provides lower levels of every public good and has a higher price of housing than some feasible alternative will be dominated by that alternative. The decision to live in a dominated community can only be explained by the presence of unobserved public goods. Thus, the proportion of dominated communities isolates the importance of accounting for unobserved public goods in the estimation. Put differently, if no community is dominated then there must be some supermodular utility function which is capable of explaining every observed location choice.

Assigning a parametric form to the (supermodular) utility function, imposes additional curvature restrictions on the relationship between the price of housing and $\bar{\xi}$.⁷⁰ Consequently, even if no community is dominated, the model may be unable to explain observed choices. In other words, satisfying the curvature restrictions may require some additional smoothing during the kernel estimation of $\bar{\xi}$. More formally, let b^* denote the minimum bandwidth on the kernel estimator such that no community is dominated, and let b^{**} denote the minimum bandwidth that allows the model to explain every observed choice. With a supermodular utility function, such as the parameterization in (5.1), it must be the case that: $b^* \leq b^{**}$. Therefore, the difference between the value for $\bar{\xi}_j$ estimated using b^* and the value for $\bar{\xi}_j$ estimated using b^{**} provides an indication of the extent to which $\bar{\xi}_j$ must be adjusted to satisfy implicit curvature restrictions imposed by the parameterization for utility.

In the current application, dominated communities and curvature restrictions both influenced the estimates for $\bar{\xi}$. Overall, 88 of the 122 communities in the San Francisco-Sacramento region are dominated. For example, table 5.5 illustrates how the Hayward Unified school district is dominated by Vallejo City, Windsor, and Benicia. The large share

⁶⁹ In this context, supermodularity requires utility to be increasing over increasing bundles of public goods and $1/p$, where bundles are ranked according to the usual vector ordering. More generally, supermodularity restricts the complementarity between the elements of a set. See Topkis (1998) for a formal definition.

⁷⁰ Feenstra and Levinsohn (1995) and Epple and Sieg (2005) also observe that the shape of the utility function imposes curvature restrictions on $\bar{\xi}$.

of dominated communities underscores the importance of accounting for unobserved public goods in the estimation.

The strategy for estimating (5.6) was to start with a guess for the bandwidth and then increase it incrementally until every observed choice could be explained⁷¹. Initial bandwidths were chosen under the assumption that the underlying distributions of price, air quality, and school quality are each normal. The resulting estimates for $\bar{\xi}$ reduced the number of dominated communities from 88 to 28. Then the bandwidths were increased proportionally and the estimation repeated until no community was dominated. Not surprisingly, even with no dominated communities, the estimator was still unable to find values for the preference parameters capable of rationalizing the choice to live in 61 of the communities. This demonstrates the importance of adjusting $\bar{\xi}$ to account for curvature restrictions. The additional “smoothing” required to rationalize these 61 choices moved the average community 5.5 places in the ranking by $\bar{\xi}$.

Table 5.6 summarizes the resulting distribution of values for $\bar{\xi}$ by reporting the highest, median, and lowest ranks for the communities within each PMSA. In general, unobserved public goods become increasingly important in explaining location choices as one moves closer to the San Francisco Bay. For example, the lowest ranking community in the San Francisco PMSA (55) is ranked above the highest ranking community in Santa Rosa (59), Yolo (82), and Sacramento (92). Table 5.7 provides a specific example of this trend by adding the estimates for $\bar{\xi}$ to the Hayward example discussed earlier. To explain why households would choose to live in Hayward, the estimator had to assign it a higher value for $\bar{\xi}$ than the other three communities. The relatively high value may reflect Hayward’s proximity to San Francisco. Hayward is located on the eastern shore of the San Francisco Bay, a 25 mile drive to the city of San Francisco. In comparison, Windsor is located about 60 miles north of San Francisco, and Vallejo City and Benicia are both located about 40

⁷¹ This process used a multiplicative Gaussian kernel, following Matzkin (2003). Also, following Bajari and Benkard (2005), $\bar{\xi}$ was rescaled by taking logs and normalizing the lowest value to be positive. This monotonic transformation simply introduces some additional smoothness into the relationship between prices and $\bar{\xi}$.

miles northeast. Since Vallejo City and Benicia are adjacent, the observation that they were assigned approximately the same value for $\bar{\xi}$ is consistent with the idea that $\bar{\xi}$ is mainly reflecting a community's location relative to the San Francisco Bay Area.

Overall, the distribution of values for $\bar{\xi}$ is similar to the vertically differentiated model. The average community differs by 7 places in the ranking by $\bar{\xi}$ between the two models. The main difference between the two distributions is that the horizontal one depicts a closer relationship between $\bar{\xi}$ and the price of housing. In the horizontal (vertical) case, the average community's price rank differs by 2 (8) places from its ranking by $\bar{\xi}$. This is not surprising given that the horizontal model identifies $\bar{\xi}$ directly from price variation whereas the vertical model defines $\bar{\xi}$ as the residual to a moment condition.

Some of the unobserved public goods that seem likely to be influencing the spatial pattern of $\bar{\xi}$ include climate, open space, and cultural amenities. The San Francisco Bay Area generally has the mildest weather in the study region. It also has the most opportunities for dining and nightlife. For example, San Francisco has more than twice as many entries in Zagat's online guide to restaurants and nightlife as the next most reviewed city in the state (Los Angeles)⁷². Finally, the Bay Area has a relatively large share of land in open space. The last two columns in table 5.6 show the number of state parks in each PMSA along with the share of land in state parks. The San Francisco, San Jose, and Santa Cruz PMSAs have the highest median values for $\bar{\xi}$ and the largest share of land in state parks. This pattern is consistent with previous sorting applications which have found open space to be an important determinant of where households locate (Walsh [2003]).

(D) *Partitioning Preference Space* $A(\alpha, \gamma, \theta)$

The partitioning process was performed in two stages. First, during the iterative part of the estimation, the single-market version of the model was used to recover $[\alpha, \gamma]$ from observed

⁷²As of July, 2006, zagat.com lists 1025 entries for dining and nightlife in San Francisco compared to 389 for Los Angeles. The city of Sacramento was not listed.

sorting across the 122 house locations. Sampling uniformly over this partition according to the population of retired households produced the joint distribution of income and preferences that was used to estimate ρ and $\bar{\xi}$. These estimates were then used in the dual-market version of the model to recover $[\alpha, \gamma, \theta]$ from observed sorting across the 268 job-house locations.

Recall from chapter 2 that the logic of revealed preferences may not fully bound regions that correspond to locations with extreme provision of public goods. In the partition shown in figure 5.4A for example, region C does not have an upper bound in the α dimension. To address this problem, absolute upper and lower bounds had to be imposed on each dimension during the Gibbs sampling process. Moreover, since the sampling process draws a finite number of points from each region of the partition, it is important to set bounds that ensure the algorithm will sample over the economically relevant portion of the unbounded regions. To illustrate this issue, equation (5.7) expresses the MWTP for air quality as a function of the heterogeneous preference parameters.

$$(5.7) \quad MWTP(g_{air,z}) = \frac{\partial V / \partial g_{air,z}}{\partial V / \partial Y} = \frac{y_{i,z}^v \cdot \alpha_i \cdot \bar{g}_{i,z}^{\rho-1} \cdot \gamma_{i,air}}{V_i^\rho - \alpha_i \cdot \bar{g}_{i,z}^\rho}, \quad \text{where}$$

$$\bar{g}_{i,z} = \gamma_{i,air} air_z + \gamma_{i,school} school_z + \gamma_{i,\xi} \bar{\xi}_z, \quad \text{and} \quad y_{i,s,z} = \hat{y}_i + \theta_1 w_{s,z} (1 - \theta_{i,2} t_{s,z}).$$

All else held constant, the MWTP approaches infinity as α increases. Therefore, regions of the partition that do not have an upper bound on α will not have an upper bound on the MWTP. For region C of the partition in figure 5.4, setting the upper bound on α to an arbitrarily large value (e.g. 1000) may produce a set of draws where most of the points translate into measures for the MWTP that exceed income. To avoid this outcome, “natural” bounds were placed on θ_1 , θ_2 , and the γ 's, allowing the bounds for α to be set based on prior assumptions about the range of plausible values for the MWTP.

The job skill parameter (θ_1) was bounded by 0 and 1.5. The lower bound implies the worker's idiosyncratic skills prevent them from gaining employment in any location other than their current niche, whereas the upper bound implies the worker is overqualified and

could make 150% of the market wage in alternative job locations. θ_2 was bounded by 0 and 1, allowing a worker's opportunity cost of time to range from 0 to their wage rate. Finally, the weights in the public goods index are normalized to sum to 1, allowing α to represent the magnitude of preferences. Its lower bound was set to 0, restricting MWTP for public goods to be nonnegative. Its upper bound was set to correspond to a \$500 MWTP for improved air quality.

More precisely, the upper bound on α sets a \$500 limit on an individual household's willingness-to-pay for a 1 part per billion (ppb) reduction in the annual average of the top 30 1-hour daily maximum readings for ozone concentrations. This measure is not directly comparable with estimates for the MWTP in much of the existing literature where air quality is typically measured by particulate matter or by the number of days during a year that ozone levels exceed state or federal standards. However, to the extent that all of these measures are simply different proxies for clean air, they can be compared in terms of a common proportionate change. Sieg et al. (2004) use this logic to translate the range of estimates in the existing literature into measures that would be comparable to the MWTP for a 1.5 ppb reduction in ozone concentrations. Converted to year 2000 dollars, the range is \$11 to \$231. Measured in these normalized units, the upper bound on α would imply a MWTP of \$750.

(i) Single-Market Results

In the single-market version of the model, the approximation to the partition of preference space is defined by 1,220,000 points—1000 points drawn from each of the 122 regions at 10 different levels of income. The 10 income “nodes” were defined as the midpoints of the income bins used to classify households in the census data. The resulting partition exhibits the same revealed preference logic as the 4-community example that was discussed in detail in the previous chapters (i.e. figure 5.4). The main difference is that adding $\bar{\xi}$ to the public goods index requires adding γ_{ξ} as a fourth dimension of preference space.

For communities that are strictly ordered according to their bundle of public goods, the partition displays the same pattern of stratification as the vertical differentiated model.

For example, figure 5.6A shows the preference regions in α, y space for four communities that follow a strict vector ordering according to prices and public good provision: Piedmont, San Mateo, Shoreline, and Esparto. Piedmont provides more of every public good than San Mateo, which provides more of every public good than Shoreline, and so on. Thus, every household will perceive Piedmont as providing a higher quality bundle of public goods than San Mateo regardless of the relative importance they assign to air quality, school quality, and $\bar{\xi}$.⁷³ Given the unanimous ordering by public good provision and the identical price ordering, a household's choice to live in Piedmont reveals that they have stronger preferences for public goods relative to private goods compared to households with the same income in San Mateo. The same is true for income given preferences. The resulting pattern of stratification shown in panel *A* of figure 5.6 is the same as in the vertically differentiated model (see figure 2.7). Panel *B* provides a different view of the same preference regions. Measuring $\ln(\alpha)$ on the z -axis and conditioning on income, the figure shows that the approximations to the four preference regions are ordered such that communities with more public goods have “higher” preference regions.

A second general feature of the partition is that households who have chosen to live in communities with *relatively* high values for a particular public good tend to be assigned *relatively* high values for the index weight corresponding to that public good. This is illustrated in figure 5.7. Panel *A* displays the (normalized) levels of air quality and school quality for communities throughout the study region. Of the three school districts that are highlighted in the figure, Sunol Glen and Milpitas provide more of every public good than Pittsburg. Therefore, their preference regions are located above Pittsburg's in the α dimension, as shown in panel *B*. However, notice that households in Sunol Glen and Milpitas have overlapping ranges of values for α . This occurs because the two communities are not strictly ordered by their provision of public goods. Sunol Glen has higher quality schools and Milpitas has cleaner air. Otherwise, they are nearly identical; the price of housing and provision of $\bar{\xi}$ differ by approximately 1% between the two communities.

⁷³ This will be true for any indirect utility function that is supermodular in public goods and $1/p$.

Thus, the choice between Sunol Glen and Milpitas helps to identify households' preferences for air quality relative to school quality. This logic underlies the “clean” result in panel *B* that households in Sunol Glen have strictly higher relative preferences for school quality. Similarly, the ratio of $\gamma_{air}/\gamma_{school}$ is strictly higher for households in Pittsburg, which offers the highest ratio of air quality to school quality out of the three communities.

The size and shape of each preference region reflects the substitution possibilities available to the households in the corresponding community. In general, preferences are better identified for households that live in communities with closer substitutes. This can be seen by comparing the projections of the three preference regions onto $\gamma_{air} > \gamma_{school}$ space in figure 5.7B with the locations of their communities in figure 5.7A. There are at least four or five other communities that are very similar to Milpitas in their provision of air and school quality. Consequently, Milpitas has a relatively small preference region compared to Pittsburg, which has no close neighbors. Of the three communities, Sunol Glen provides the best illustration of how substitution possibilities identify relative preferences. Its preference region is relatively narrow in the γ_{air} dimension because there are two other communities with nearly the same school quality, one with slightly clearer air and one with slightly dirtier air. Because the reverse is not true, its preference region is relatively wide in the γ_{school} dimension. Notice that Sunol Glen's oblong preference region is approximately the right shape to “fill” the empty space surrounding Sunol Glen in public goods space. This is also true for the other two communities.

Finally, although the Gibbs algorithm sampled uniformly over each preference region, there is considerable sparseness near some of the edges. For example, in $\gamma_{air} > \gamma_{school}$ space, there are relatively few points in the upper left corner of the region for Pittsburg and also in the right corner of the region for Sunol Glen. In both cases, the preference regions are pyramid shaped and the sparseness occurs in the tip which would be consistent with a constant density of points. This appears to be a general feature of the single and dual-market versions of the partition.

(ii) Dual-Market Results

In the dual-market version of the model, the approximation to the partition of preference space is defined by 58,960,000 points—1000 points drawn from each of the 268 regions for each of the 220 (occupation, non-wage income) pairs⁷⁴. The main difference from the single-market partition is that adding work destinations to the choice set expands the borders of the preference sets. Intuitively, heterogeneity in job skill and the opportunity cost of time provide new ways to explain observed location choices.

Figure 5.8 provides representative examples of how the preference regions differ between the single and dual-market partitions. Panels *A*, *B*, and *C* project the preference regions recovered for architects and engineers living in the Acalanes school district onto $\gamma_{air}, \gamma_{school}$ space. In the single-market case (panel *A*) the choice to live in Acalanes reveals strong preferences for school quality relative to air quality because Acalanes has high quality schools (90th percentile) and low quality air (14th percentile). Of all the possible job destinations for architects and engineers who live there, the Oakland PMSA requires the shortest commute. Therefore, the choice to live in Oakland may reveal a high opportunity cost of time rather than strong preferences for school quality. This possibility is reflected in the way the preference region in panel *B* is “stretched” to the left compared to panel *A*. The lowest values for γ_{school} correspond to high values for the opportunity cost of time parameter (θ_2). In contrast, the preference region is stretched to the right for workers who make the relatively long commute to San Francisco. In this case, the highest values for γ_{school} are paired with low values for the job mobility parameter (θ_1). For an architect or engineer who is “stuck” working in San Francisco, the choice to live far from their job reveals strong preferences for the public goods provided by that community—in this case school quality.

Panels *D*, *E*, and *F* of the figure display the preference regions for managers who live in the Berryessa school district. Berryessa has average values for air and school quality and the households who live there are assigned similar values for γ_{school} and γ_{air} in the single-

⁷⁴ This process followed a burn-in of 100 draws to reduce sensitivity to starting values. The sampling process took approximately 20 hours on a standard desktop computer.

market case; i.e. $\gamma_{air}/\gamma_{school} \approx 1$. Panel *E* illustrates that adding job opportunities to the model can lead to a much higher ratio for managers who work in the San Jose PMSA. The reason is that San Jose is the highest paying job destination and Berryessa offers one of the shortest commutes. This explains why managers with relatively strong preferences for air quality would choose to live there. Despite the high wages in San Jose, some households in Berryessa make the seemingly irrational choice to commute a longer distance to the Oakland PMSA where managers are paid \$13,000 less. As chapter 3 discussed, this choice can be explained by a low value for θ_1 . However, a low value for θ_1 also implies that the manager will be paid below the market wage if they were to move to a community that is not within commuting distance of their current job. The implied drop in income decreases the attractiveness of some of the communities that had neighboring preference regions in the single market version of the model. As a result, the preference region in panel *F* is stretched to the left compared to panel *D*.

While the dual-market specification uses more information to identify preferences, more is required of this information; i.e. two additional dimensions of preference space (θ_1, θ_2) must be partitioned. The Acalanes and Berryessa examples demonstrate how this expands the range of values for the weights in the public goods index that are consistent with the choice of each community. In other words, recognizing that job opportunities influence sorting behavior implies greater diversity in relative preferences for public goods within communities. This effect can help to explain Rhode and Strumpf's (2004) observation that communities in the United States have become more alike as mobility costs have declined over the past 150 years.

If local public goods are the only factor affecting location choices, a decrease in moving costs would make it easier for households to stratify according to their relative preferences, decreasing the amount of preference heterogeneity within communities. As the diversity in preferences within communities decreases, one would expect them to become more specialized in their provision of local public goods. However, Rhode and Strumpf find just the opposite—as mobility costs have declined over time so has the between-community variation in proxy measures for public good provision such as local taxes per capita and

school taxes per capita. To illustrate how labor markets can help to explain this outcome, figure 5.9 contrasts the single and dual-market patterns of preference stratification for managers living in the Pittsburg, Milpitas, and Sunol Glen school districts. Adding job choices to the model creates substantial overlap in the preference regions even though households are assumed to be freely mobile. An alternative interpretation is that households are *not* freely mobile because specialization in job skill and the shadow value of time make it costly for them to sort solely according to their preferences for public goods. Therefore, if workers become more specialized over time (i.e. θ_1 decreases) their ability to sort according to their preferences for public goods will decrease, counteracting the effect of lower moving costs. This effect could help to explain the convergence in provision of public goods reported by Rhode and Strumpf.

Additional insight into the economic implications of including job opportunities in the sorting framework can be derived by analyzing the model's implications for partial equilibrium welfare measures.

VI. Quantifying the Economic Implications of the Model

The second-stage results from each of the three sorting models characterize the distribution of preferences for public goods in the population of households who live in the San Francisco-Sacramento area. However, most of the structural parameters do not have a direct economic interpretation. Therefore, to compare the implications of the three models, the information about preferences was translated into the MWTP for an air quality improvement. For the vertical model, this simply requires drawing a sample of households from the joint distribution of income and preferences defined by the parameter estimates for $\mu^\alpha, \mu^y, \sigma^\alpha, \sigma^y, \lambda$ and using equation (5.7) to convert these draws into measures for the MWTP. In contrast, using the results from the two horizontal models to generate a distribution of MWTP requires drawing a sample of households from the partitions of preference space. Part (A) describes how the sampling process was undertaken and part (B)

analyzes how the resulting distributions of MWTP vary across space, income, and occupations.

(A) *Generating a Distribution of Preferences for the Population of Households*

Using the partition of preference space to generate a distribution of preferences requires making an assumption about the way preferences are distributed within each region of the partition. In order to evaluate the sensitivity to these distributional assumptions, bounds were constructed on the marginal distribution of each heterogeneous parameter for the population of households in the study region. For example, the “lower bound distribution” for $F(\gamma_{air})$ was constructed by assigning every household the lowest possible value of γ_{air} that is consistent with its observed location choice. The census data report 566 households with a primary earner in the managerial occupation who live in the Berryessa school district, work in the San Jose PMSA, and have total income of \$112,500. Therefore, 566 draws were taken on the lowest value for γ_{air} in the region of the partition that corresponds to this household “type”. Likewise, 566 draws were taken on the highest value for γ_{air} while constructing the upper bound distribution. This process was repeated for every household type so that the resulting distributions represent all 3.2 million households in the study region. Figure 5.10 shows results from the dual-market specification. Any assumption about the joint distribution of preferences will lead to a marginal distribution that falls within the bounds in the figure (depicted by the two solid lines). The dashed lines correspond to the assumption that preferences are joint-uniformly distributed. That is, they were constructed by sampling uniformly over each region of the partition according to the population of households in the corresponding community.

In figure 5.10, the marginal distributions for α , γ_{air} , γ_{school} , and γ_{ξ} have relatively narrow bounds over most of their support and relatively wide bounds for a small share of households. The job skill and shadow value of time parameters have less informative bounds. Recall from chapter 3 that θ_1 is not bounded from below by the logic of revealed preferences. The vertical line at $\theta_1 = 0$ simply reflects the maintained assumption that job

skill cannot be negative. For θ_2 , the maintained assumption that the opportunity cost of time cannot exceed the wage rate (i.e. $\theta_2 \leq 1$) was binding for approximately 50% of households. In other words, the location choices observed for these households would also be consistent with an opportunity cost of time that exceeds their wage rate. This reflects the fact that most households work in the PMSA closest to their home. The underlying revealed preference logic does not provide an upper bound on their opportunity cost of time for households who have chosen the shortest possible commute.

Figure 5.11 contrasts the distributions for γ_{air} , γ_{school} , and γ_{ξ} calculated from the single and dual-market partitions. Not surprisingly, the single-market distributions are better identified in the sense that there is less difference between the upper and lower bounds. Thus, distributional assumptions will play a larger role in identifying preferences in the dual-market case. To compare the shape of the distributions generated from the two partitions, table 5.8 reports means and standard deviations for the uniform distribution of each preference parameter (i.e. distributions denoted by dashed lines in the figure). In the dual-market case, the means for α , γ_{air} and γ_{school} are all slightly larger and the mean for γ_{ξ} is slightly smaller. Intuitively, without job opportunities to help explain location choices, the single-market version of the model has to assign more importance to unobserved public goods to rationalize observed behavior. The larger standard deviations on γ_{air} , γ_{school} , and γ_{ξ} in the dual-market case reflect the way that job opportunities tend to widen the bounds on the preference regions.

Table 5.8 also reports the point estimates for α and γ_{air} from the vertically differentiated model. They are not comparable to the horizontal results in terms of magnitude since they correspond to different estimates for ρ and $\bar{\xi}$. Nevertheless, there is a striking difference between the relative values for the (average) weights estimated for the horizontal model and the (constant) weights estimated for the vertical model. The ratio of γ_{air} to γ_{school} in the two horizontal models is seven to ten times larger than in the vertical case. This could be due to the many differences between the two estimators, or it could

simply reflect the large standard error on the vertical point estimate for γ_{air} .

Finally, of all the heterogeneous preference parameters, θ_2 has the most straightforward interpretation. Under the uniform distributional assumption, its mean value of 0.401 implies the mean shadow value of time is approximately 40% of the wage rate. This is quite similar to the rule-of-thumb (33%) that is often used in recreation demand studies (Phaneuf and Smith [2005]). However, figure 5.10 illustrates that the mean value for θ_2 will depend largely on assumptions about the shape of its distribution.

(B) Marginal Willingness-to-pay for an Air Quality Improvement

The partitions were translated into distributions of the willingness-to-pay for a marginal (1 ppb) reduction in ozone concentrations by sampling from each region of preference space according to the associated population of households and then converting each draw into the measure of MWTP in equation (5.6). This approach was used to generate three distributions, each corresponding to a different assumption about the way preferences are distributed. First, the joint uniform distribution depicted by the dashed lines in figures 5.10 and 5.11 was translated into a distribution of MWTP. Then, upper and lower bounds on that distribution were generated in a parallel fashion to the bounds on the individual preference parameters. That is, an upper (lower) bound distribution was produced by assigning every household the highest (lowest) possible MWTP consistent with its observed location choice. This process was repeated for both the single and dual-market partitions.

The difference between the upper and lower bound distributions can be used to measure the economic significance of assumptions on the distribution of preferences. Table 5.9 reports the share of households within 7 different “identification intervals”. For example, the difference between the highest and lowest MWTP that would be consistent with observed location choices lies between \$0 and \$10 for 4.5% of households in the single-market case. In other words, the MWTP is identified to within \$10 for these households. Likewise, the MWTP is identified to within \$25 for 18.6% of households (14.1% + 4.5%). Moving from the single to the dual-market case decreases the share of households for whom the MWTP is precisely estimated. This is consistent with the observation that the dual-market preference

regions typically have wider bounds. The range of consistent measures exceeds \$250 for 32.3% of households. For them, MWTP will be identified almost entirely by distributional assumptions.

Table 5.10 provides summary measures of the MWTP distributions and compares them to the corresponding results from the vertical model. The top row reports the average per/household MWTP for all households in the study region. For example, when every household is assigned the minimum (maximum) MWTP that would be consistent with its observed location choice in the dual-market version of the model, the average per/household MWTP in the population of households is \$33 (\$226). Notice the range of estimates in the dual market case (\$33 to \$226) contains the range in the single-market case (\$57 to \$168) which contains the point estimate from the vertical model (\$83). This illustrates the economic relevance of the “bias/variance” tradeoff described in chapter 2. That is, if the depiction of utility in the dual-market case represents the “truth”, then treating income as exogenous and preferences as vertically differentiated will have two effects. It will bias the resulting welfare measures and it will decrease the sensitivity of those measures to assumptions on the distribution of heterogeneous preference parameters.

The bottom three rows in table 5.10 compare the average per/household MWTP across individual communities. Results are reported for the lowest, median, and highest communities. For example, the median community in the vertical model has an average MWTP of \$50 compared to \$81 and \$97 in the two horizontal models (under the uniform assumption). This illustrates another general feature of the results: conditional on the uniform assumption, introducing horizontal differentiation and accounting for job opportunities both tend to increase the MWTP. One exception is San Francisco’s second supervisorial district. As the most expensive community, the vertical model assigns households who live there the highest values for $\alpha | y$ from the right tail of the (lognormal) distribution. The resulting average MWTP of \$819 exceeds the upper bound of \$500 that was imposed on the two horizontal models. In the single-market case, this upper bound truncated the preference regions for approximately 6.0% of households, compared to 13.8% in the dual-market case. To further analyze the properties of the MWTP distributions, the

remainder of this section focuses on the results generated under the uniform assumption.

Figure 5.12 plots the average per/households MWTP in each community against the level of air quality in that community. All three sorting models depict a similar increasing trend. Intuitively, all else held constant, the choices to live in a community with cleaner air reveals a higher MWTP for air quality. Moving from vertical to horizontal differentiation and introducing labor markets increases the variability in this relationship by offering alternative explanations for why households with strong preferences for air quality would choose to live in communities with high ozone concentrations. For example, the average MWTP in Los Altos is \$80 in the vertical model compared to \$251 in the (single-market) horizontal model. While Los Altos has relatively low air quality, it has the highest school quality in the study region. The horizontal model recognizes that households with strong preferences for air quality may choose to live there because they have even stronger preferences for school quality. Similarly, adding job opportunities to the model increases the MWTP in Natomas from \$163 to \$245, acknowledging that households with strong preferences for air quality may choose to live there because it offers the shortest commute to job destinations in the Sacramento PMSA.

A second reason why the MWTP for air quality tends to increase in the level of air quality is that households with higher incomes tend to live in communities with cleaner air. Figure 5.13 graphs the relationship between air quality and average household income across communities. A slight upward trend is visible. To illustrate the relationship more directly, figure 5.14 displays the average MWTP for households in each of the 10 income bins in the Census data. Since the expression for the MWTP in (5.7) is increasing in income, it comes as no surprise that average MWTP is increasing in average income. For the first nine income bins, the dual-market estimator produces the highest estimate for the average MWTP followed by the single-market horizontal and vertical models. The reversal in the last bin simply reflects the way that the two horizontal models truncate the right tails of the income and preference distributions.

Table 5.11 provides a different view of the relationship between air quality and income by reporting the average MWTP by occupation and work destination. Households

with primary earners working in the Santa Cruz PMSA have the highest MWTP, followed by those with primary earners who work in Santa Rosa. This pattern can be explained by four observations. First, approximately three quarters of the people who work in Santa Cruz and Santa Rosa also live there (table 4.2). Second, workers in every occupation are paid less in those two PMSAs than in most or all other job destinations (table 4.5). Third, after the San Francisco PMSA, Santa Cruz and Santa Rosa have the highest overall air quality (table 5.11, bottom row). Fourth, housing is much more expensive in communities near the San Francisco PMSA because they tend to have the highest values for $\bar{\xi}$ (tables 5.2 and 5.6). Together, these observations suggest that living and working in Santa Cruz or Santa Rosa will be the best choice for households with relatively strong preferences for air quality (i.e. high $\gamma_{air}/\gamma_{\xi}$). Changing jobs could lead to a higher salary but, for most households, would also require moving to a different community with lower air quality. By forgoing the opportunity to earn higher wages, households in Santa Cruz and Santa Rosa reveal a relatively high MWTP for reduced ozone concentrations.

After Santa Rosa and Santa Cruz, households who work in the Yolo and San Francisco PMSAs tend to have the next highest average MWTP. For San Francisco, this reflects the high air quality and high average incomes in the communities surrounding the San Francisco Bay. For Yolo, the relatively high average MWTP is counterintuitive since most of the nearby communities have low air quality. However, air quality in these communities is high relative to the estimated values for $\bar{\xi}$. Moreover, Yolo offers the shortest commute times. Therefore, households with a high opportunity cost of time and relatively strong preferences for air quality may choose to live there. The Oakland, San Jose, and Vallejo PMSAs are centrally located in the study region and the households who work there live in communities with a wide range of values for air quality. This helps to explain why they tend to have similar values for the MWTP. Households who work in the Sacramento PMSA generally have the lowest MWTP because the nearby communities have the lowest air quality in the study region.

Finally, compared to the results from previous studies, the dual-market estimates for the MWTP are relatively high. Converting the range of normalized values for the existing

literature (discussed in section V) into measures that would be equivalent to the average MWTP for a 1 ppb ozone reduction implies a range from \$7 to \$154 (year 2000 dollars). The higher range produced by the dual-market estimator (\$33 to \$226) could stem from methodological differences or simply from differences in the study region. The \$7 and \$154 estimates are both for Los Angeles which has much higher ozone concentrations than the San Francisco-Sacramento area. Moreover, median income in the San Francisco CMSA is 35% higher than in the Los Angeles CMSA. If Northern and Southern California were considered as part of the same choice set, the relationship between MWTP, air quality, and income displayed in figures 5.12 and 5.13 would imply that households in San Francisco and Sacramento would tend to have a higher MWTP than those in Los Angeles.

From a methodological perspective, the best comparison to the existing literature is to Sieg et al's (2004) application of the single-market vertical model to Los Angeles in 1990. They report an average MWTP of \$66. However, the average level of ozone concentrations across the communities in their application is 150 ppb, compared to a maximum of 109 here. The Sacramento PMSA provides the closest approximation to the income and ozone conditions in Los Angeles. The average level of ozone concentrations for the communities physically located in Sacramento is 94 ppb and the median income is 1.5% higher than in Los Angeles. For the households who live in these communities, the average MWTP predicted by the single-market vertical model is \$23, compared to \$68 and \$73 for the two horizontally differentiated models (under the uniform assumption). The low estimate for the vertical model reflects the fact that the communities in the Sacramento PMSA have the lowest housing prices in the study region (table 5.2). Therefore, conditional on income, they are assigned the lowest values for α , which imply the lowest values for the MWTP. The horizontal models also assign relatively low values to households in these communities, but recognize that variation in relative preferences and job opportunities may induce some households with relatively strong preferences for air quality to locate there.

VII. Conclusions

This chapter presented the results from using three versions of the empirical model to estimate households' preferences for public goods in the San Francisco-Sacramento region of California. The process began by calculating an index of prices for the bundle of public goods provided by each of the 122 communities. This was used together with community-specific distributions of income and rents to estimate the demand for housing. Both the demand for housing and the price index were then treated as known constants during the second stage of the estimation which used a parameterization of the indirect utility function and households' observed location choices to simultaneously recover: (1) the elasticity of substitution between public and private goods; (2) a composite unobserved public good for each community; and (3) a set of values for the heterogeneous preference parameters capable of explaining observed location choices. Translating these sets into a distribution of preferences for the population of households required choosing a shape for that distribution. The shape assumption was found to have a substantial impact on estimates of the average MWTP for air quality. In the dual-market case, estimates for the average MWTP ranged from \$33 to \$226.

Moving from the Epple-Sieg model to the dual-market framework increased the average MWTP by 47% under the "naïve" assumption that preferences are distributed uniformly within each region of the partition of preference space. Under this assumption, relaxing vertical differentiation and introducing labor market choices both increased the average MWTP. More generally, these extensions also increased the extent to which the average MWTP depends on maintained assumptions about the distribution of heterogeneous preference parameters in the population of households.

Recovering an approximation to the partition of preference space provided a new perspective on the partial equilibrium welfare measures generated from the model. However, for the unbounded regions of the partition it also required selecting an arbitrary upper bound on the MWTP. The choice for the upper bound may have a large impact on measures of the MWTP for households who live in the corresponding communities. Future work could consider alternative strategies for bounding the partition. Chapter 7 suggests some

possibilities.

The labor market side of the model provides a second opportunity for future research. With few work destinations (8) compared to the number of housing communities (122) there was relatively little information to identify the job skill and opportunity cost of time parameters compared to the parameters that represent preferences for public goods. In future work, this could be addressed by using a hedonic wage regression to predict the wage each worker would earn in alternative job locations based on their demographic characteristics, as in Bayer, Keohane and Timmins (2006). This approach would be enhanced by micro data on the characteristics of workers and their occupations (e.g. education, age, industry, job title). Information on specific job titles could also help to better identify job skill within the current estimation framework. Potential sources for these data are described in chapter 7.

Finally, while measures of the MWTP are useful for deriving intuition for the economic implications of the model, they fail to capture the welfare effects from general equilibrium adjustment, limiting their applicability to practical policy analysis. That is, a sufficiently large ozone reduction may lead to changes in housing prices and wage rates as some households move between communities and/or jobs. These changes will alter the welfare implications of the air quality improvement. Chapter 6 uses the estimation results to simulate how households and markets would adjust to the large scale ozone reductions required to meet California's recently revised standards for ambient ozone concentrations.

Table 5.1: Hedonic Estimation of Housing Price Indices*

Variable	Coefficient	Standard Error
Bathrooms	0.0312	(0.0017)
Bedrooms	-0.1128	(0.0019)
Building (sqft)	-0.3288	(0.0378)
Lot (sqft)	-0.0686	(0.0107)
Age	0.0570	(0.0083)
Age ²	-0.0032	(0.0002)
Lot ²	-0.0112	(0.0003)
Building ²	0.0468	(0.0027)
Age x Lot	0.0124	(0.0005)
Age x Building	-0.0266	(0.0012)
Lot x Building	0.0484	(0.0014)
Condominium (dummy)	-0.1348	(0.0022)
N	540,642	
R ²	0.8143	

* Year and community fixed effects were included.

Table 5.2: Housing Price Index by Primary Metropolitan Statistical Area

PMSA	Prices for Individual Communities		
	Minimum	Median	Maximum
Sacramento	1.00	1.22	1.77
Yolo	1.31	1.53	2.08
Oakland	1.42	2.30	4.23
San Francisco	2.54	3.74	6.51
San Jose	2.11	2.93	4.85
Santa Cruz	2.36	2.66	3.18
Santa Rosa	1.65	2.03	2.39
Vallejo	1.28	1.47	2.96

Table 5.3: Housing Demand Parameters (standard errors in parenthesis)

Specification*		Demand Constant (β)	Price Elasticity (η)	Income Elasticity (ν)	R ²
Retired Households Included	OLS	82.82 (0.17)	-0.29 (0.02)	0.49 (0.02)	0.888
	IV price = f(ozone, score)	78.07 (1.19)	-0.33 (0.03)	0.50 (0.02)	0.885
	IV price = f(ozone, score) inc = f(nonwage inc)	36.22 (1.48)	-0.35 (0.03)	0.57 (0.04)	0.876
Retired Households Excluded	OLS	29.72 (1.23)	-0.33 (0.02)	0.58 (0.02)	0.888
	IV price = f(ozone, score)	27.03 (1.23)	-0.37 (0.03)	0.59 (0.02)	0.887
	IV price = f(ozone, score) inc = f(nonwage inc)	11.97 (1.57)	-0.38 (0.03)	0.66 (0.04)	0.876

* Using housing prices estimated from a log-log specification of the hedonic expenditure function.

Table 5.4: Results from GMM Estimation of Single-Market Vertical Model

Parameter	Symbol	Bounds on Starting Value Search		Starting Values	GMM Results (4 iters)	
		Lower	Upper		Estimate	Standard Error
		(1)	(2)	(3)	(4)	(5)
mean ln(income)	μ^y	10.5	11.5	11.060	11.057	1.13E-06
standard deviation ln(income)	σ^y	0.7	0.9	0.797	0.762	1.01E-08
mean ln(preferences)	μ^α	0.0	1.0	0.861	0.874	0.219
standard deviation ln(preferences)	σ^α	0.0	0.8	0.725	0.755	1.01E-08
correlation(income, preferences)	λ	-0.5	0.0	-0.478	-0.477	7.32E-09
CES substitution parameter	ρ	-0.2	0.0	-0.022	-0.022	3.10E-11
public goods in community 1	g_1	0.1	1.0	0.306	0.310	0.158
weight on air quality in public goods index	γ_{air}	0.0	1.5	0.136	0.137	0.628

*To find starting values for the GMM estimation, the Genetic algorithm evaluated the objective function, 300,000 times. GMM was performed using the genetic starting values, a simulation size of 100,000, all tolerances set to 1e-8, and the instrument specification: Z=[ones, crank2, crank3], where "crank2" denotes a 2nd order Chebychev polynomial function of community rankings by price.

Table 5.5: Hayward Unified is Dominated by Vallejo City, Windsor, and Benicia

Community	Price	Public Goods*	
		School quality	Air quality
Hayward Unified	2.07	0.86	0.95
Vallejo City Unified	1.51	0.88	1.05
Windsor Unified	1.82	0.95	0.97
Benicia Unified	1.84	1.14	1.70

* Higher values indicate higher quality.

Table 5.6: Composite Unobserved Public Good, summarized by PMSA

PMSA	Community Rank* by ξ			State Parks	
	Highest	Median	Lowest	number	% of land area
Sacramento	92	115	122	28	2.7
Vallejo	35	103	110	5	0.2
Yolo	82	96	106	1	0.0
Santa Rosa	59	86	95	13	3.6
Oakland	13	71	100	10	3.0
Santa Cruz	25	46	67	16	12.9
San Jose	5	34	76	4	10.8
San Francisco	1	18	55	28	6.2

* 1 = highest; 122 = lowest

Table 5.7: Adding $\bar{\xi}$ Explains the Choice to Live in Hayward

Community	Price	Public Goods*		ξ
		School quality	Air quality	
Hayward Unified	2.07	0.86	0.95	0.36
Vallejo City Unified	1.51	0.88	1.05	0.15
Windsor Unified	1.82	0.95	0.97	0.26
Benicia Unified	1.84	1.14	1.70	0.15

* Higher values indicate higher quality

Table 5.8: Summary Statistics for the Heterogeneous Preference Parameters *
(standard errors in parentheses)

Parameter	Single-Market Vertical	Single-Market Horizontal	Dual-Market Horizontal
$\log(\alpha)$	0.861 (0.724)	-9.022 (4.596)	-8.890 (4.566)
γ_{school}	1.000	0.105 (0.154)	0.155 (0.199)
γ_{air}	0.133 (0.628)	0.126 (0.144)	0.152 (0.183)
γ_{ξ}		0.769 (0.213)	0.693 (0.278)
θ_1			0.460 (0.275)
θ_2			0.401 (0.319)

* The means and standard deviations for the horizontal model are based on the assumption that preferences are distributed uniformly within each region of the partition

Table 5.9: Identifying MWTP for Improved Air Quality, Horizontal Models

Model	Share of Households with $ \max(\text{MWTP}) - \min(\text{MWTP}) $ in the Range:						
	\$0-\$10	\$10-\$25	\$25-\$50	\$50-\$75	\$75-\$100	\$100-\$250	\$250-\$500
Single-Market	4.5%	14.1%	23.2%	16.9%	11.4%	16.2%	13.8%
Dual-Market	3.2%	7.6%	14.2%	12.3%	8.0%	22.3%	32.3%

Table 5.10: MWTP for Improved Air Quality, 3 Sorting Models

Average per/household MWTP for:	<u>Single-market, Vertical</u>	<u>Single-market, Horizontal</u>			<u>Dual-market, Horizontal</u>		
	$f(\alpha, y) \sim \text{joint lognormal}$	min MWTP	uniform pref.	max MWTP	min MWTP	uniform pref.	max MWTP
All Households	83	57	109	168	33	122	226
Lowest community	14	0	15	39	0	12	25
Median community	50	47	81	110	15	97	182
Highest community	819	254	396	500	253	389	500

Table 5.11: Average per/household MWTP, by Occupation and Work PMSA

Occupational Category (Standard Occupational Classification System)	Primary Metropolitan Statistical Area							
	San Jose	Oakland	Sacramento	San Francisco	Santa Cruz	Santa Rosa	Vallejo-Fairfield-Napa	Yolo
Management	188	127	99	201	281	243	137	209
Business and Financial Operations	134	91	94	181	298	257	127	197
Computer & Mathematical	187	108	102	177	275	268	132	218
Architecture and Engineering	168	96	91	178	282	261	153	204
Life, Physical and Social Science	193	126	92	190	319	272	108	216
Community and Social Service	90	88	85	158	342	259	120	171
Legal	198	137	100	225	303	260	126	233
Education, Training and Library	107	88	89	175	312	242	121	182
Arts, Design, Entertainment, Sports and Media	122	109	89	197	304	262	124	165
Healthcare Practitioners and Technicians	147	105	89	185	290	246	130	202
Healthcare Support	82	72	41	135	369	245	69	128
Protective Service	91	73	84	136	325	265	111	160
Food Preparation and Serving	68	66	55	156	346	231	76	178
Building and Grounds Cleaning & Maintenance	60	57	52	122	358	221	66	159
Personal Care and Service	76	72	55	143	359	232	70	155
Sales and Related	99	65	68	155	338	232	67	171
Office and Administrative Support	95	71	64	132	334	240	77	187
Farming, Fishing and Forestry	25	42	40	98	385	163	44	113
Construction and Excavation	102	72	76	132	325	239	108	178
Installation, Maintenance and Repair	95	65	75	127	341	250	102	171
Production	81	67	55	142	374	242	88	165
Transportation and Material Moving	69	66	52	128	371	231	74	164
Average Ozone (ppb)*	74.2	77.0	93.9	55.6	63.3	65.1	74.2	86.4

* Calculated as a weighted average over the communities physically located in each PMSA, using acreage shares for weights.

A. Maintained Assumptions on Choices and Preferences in the 3 Versions of the Model

Assumption	(1) Dual-Market Horizontal	(2) Single-Market Horizontal	(3) Single-Market Vertical
Income	Varies with job location	Fixed	Fixed
Choice set	268 (community, labor market) pairs	122 communities	122 communities
Preference heterogeneity	Horizontal	Horizontal	Vertical
Distributional assumption	None	None	$f(\alpha, y) \sim \text{lognormal}$
Restrictions on utility in (3.1)	None	$\theta_2 = 0, \theta_1 = 0$	$\theta_2 = 0, \theta_1 = 0, \gamma = \gamma_i \forall i$

B. First-Stage Strategy: Estimate Once and Use Results in all 3 Models

- i. Recover index of housing prices (p_1, \dots, p_{122}) from hedonic expenditure function
- ii. Recover housing demand parameters (β, η, ν) from housing demand function

C. Second-Stage Strategy: Estimate Vertical and Horizontal Models Independently

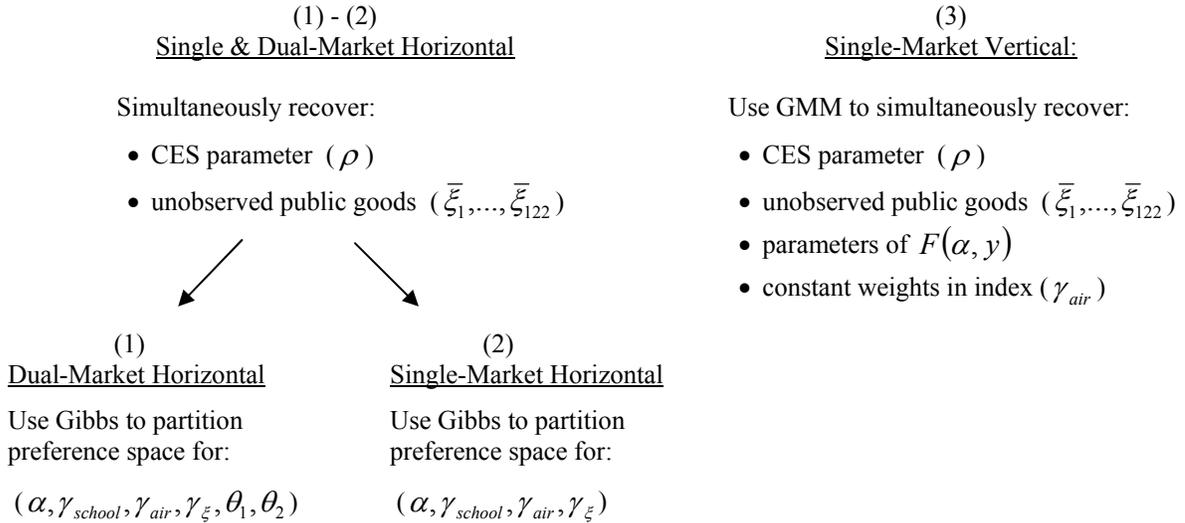


Figure 5.1: Overview of the Estimation

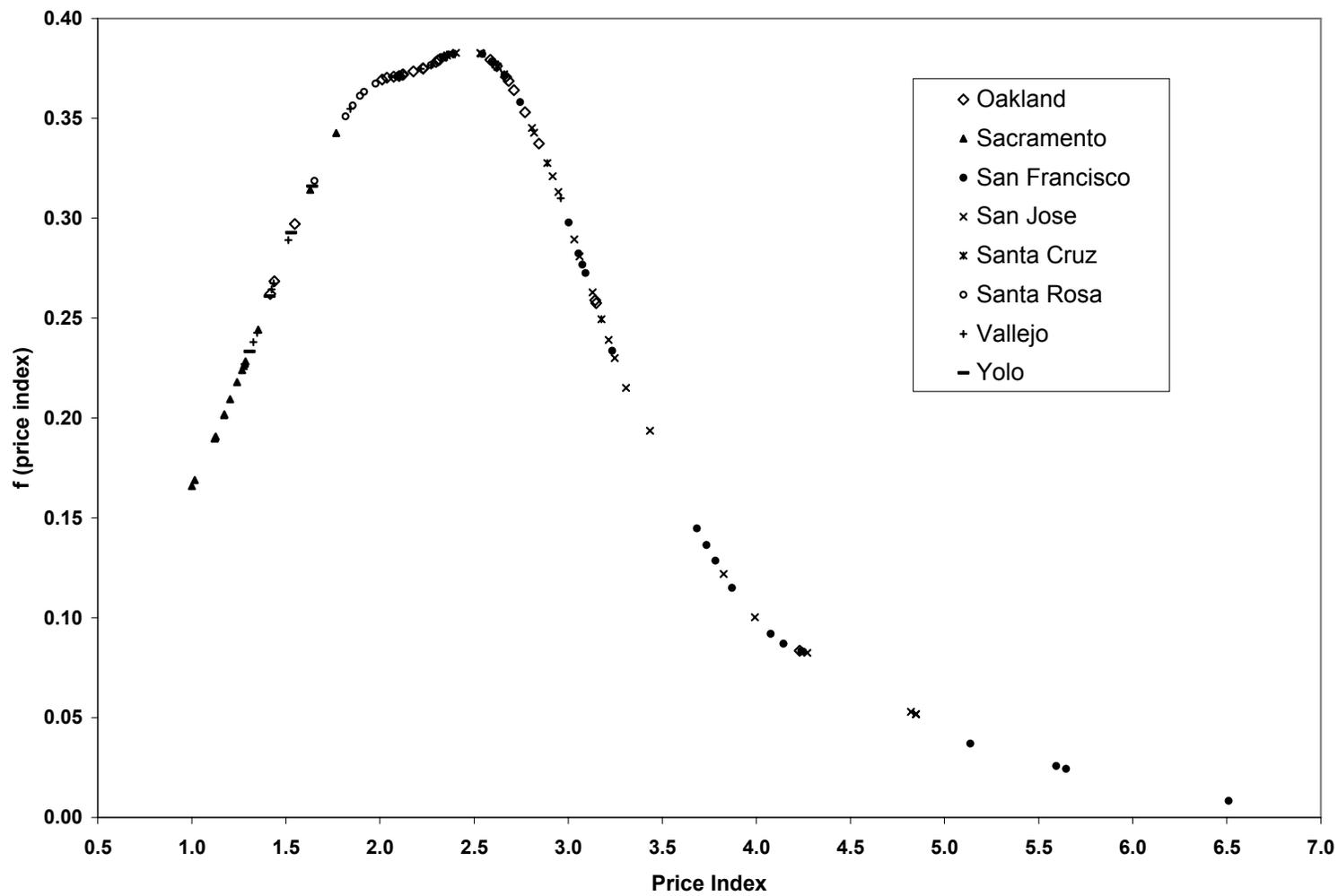


Figure 5.2: Distribution of Housing Prices (Distribution was created using a Gaussian kernel and the bandwidth that would be optimal if the underlying distribution were normal.)

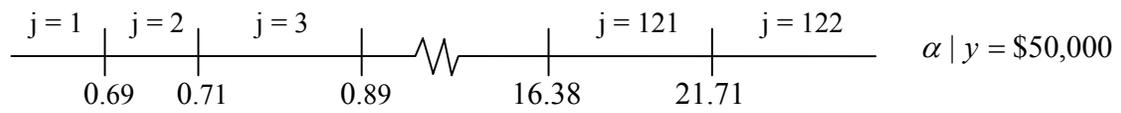
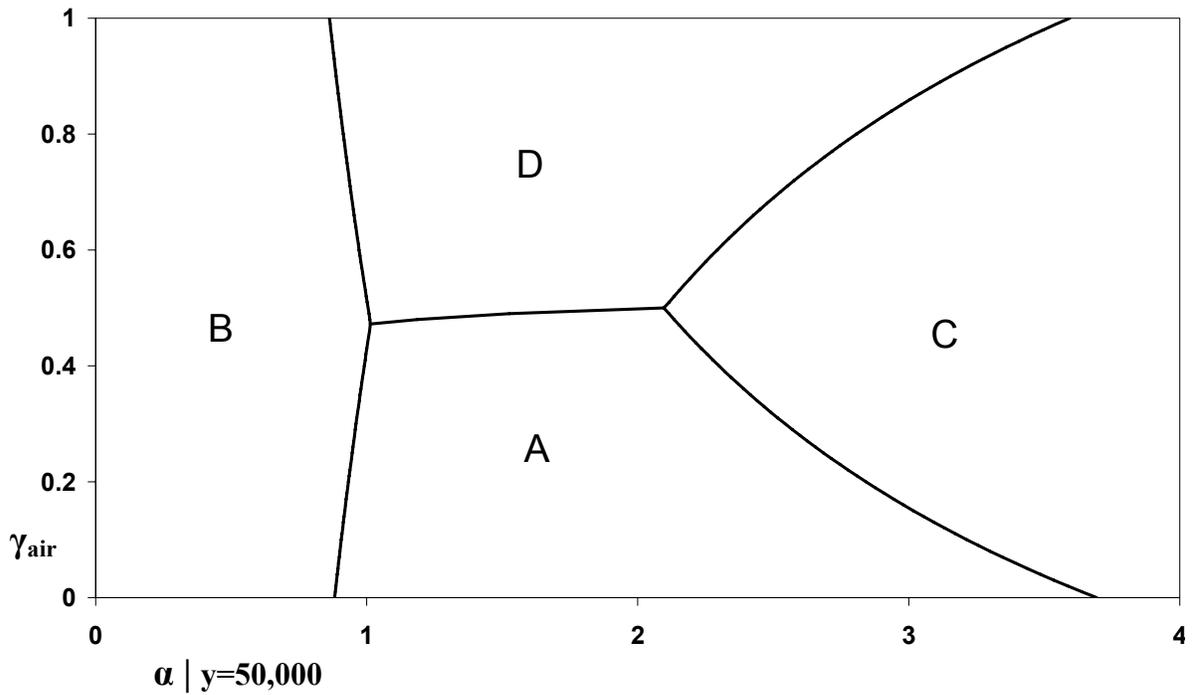


Figure 5.3: Implied Partition of Households across Communities, Given $y = \$50,000$

A. True Partition from 4-Community Example



B. Approximation to the True Partition (Different symbols represent points in each region)

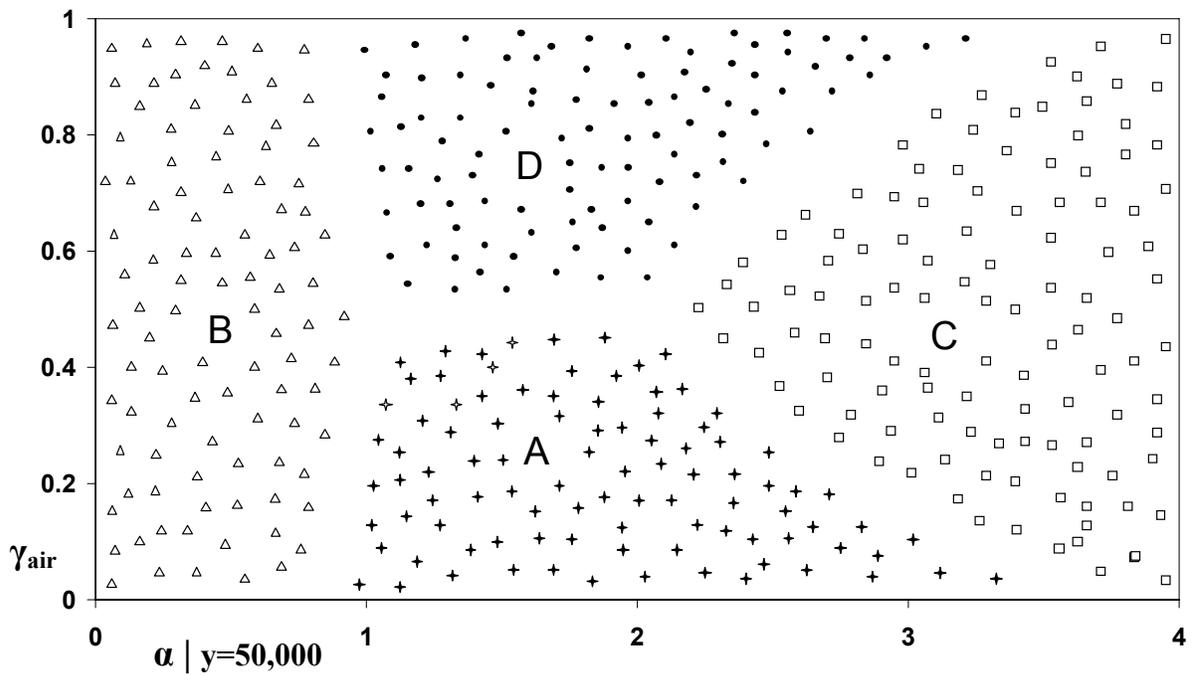
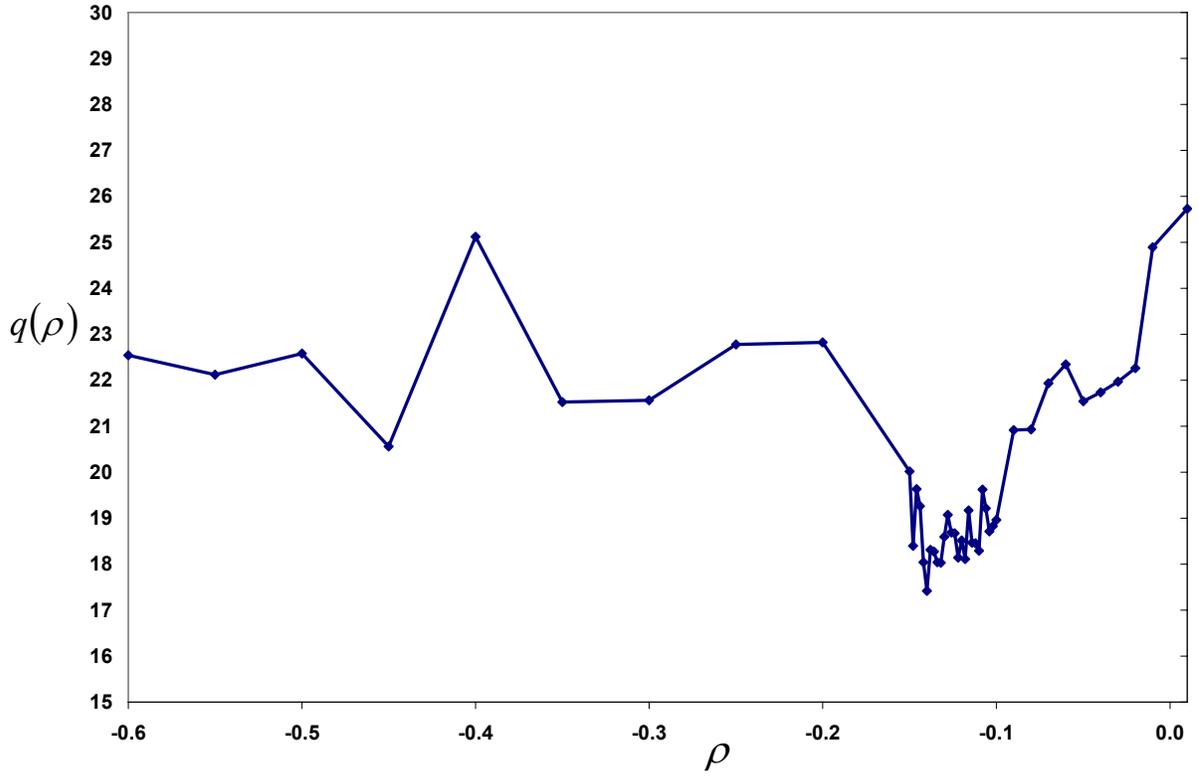


Figure 5.4: Recovering an Approximation to the Partition of Preference Space



Objective Function

$q(\rho) = q_1 + q_2 + q_3 + q_4$, where:

$$q_1(\rho) = \sum_j \sum_i \left| f_{\alpha}^i(\alpha, \gamma | y_{j,1}, \rho, \xi, \beta, \eta, \nu) - f_{\alpha}^i(\alpha, \gamma | y_{j,2}, \rho, \xi, \beta, \eta, \nu) \right|,$$

$$q_2(\rho) = \sum_j \sum_i \left| f_{\gamma_{air}}^i(\alpha, \gamma | y_{j,1}, \rho, \xi, \beta, \eta, \nu) - f_{\gamma_{air}}^i(\alpha, \gamma | y_{j,2}, \rho, \xi, \beta, \eta, \nu) \right|,$$

$$q_3(\rho) = \sum_j \sum_i \left| f_{\gamma_{school}}^i(\alpha, \gamma | y_{j,1}, \rho, \xi, \beta, \eta, \nu) - f_{\gamma_{school}}^i(\alpha, \gamma | y_{j,2}, \rho, \xi, \beta, \eta, \nu) \right|,$$

$$q_4(\rho) = \sum_j \sum_i \left| f_{\gamma_{\xi}}^i(\alpha, \gamma | y_{j,1}, \rho, \xi, \beta, \eta, \nu) - f_{\gamma_{\xi}}^i(\alpha, \gamma | y_{j,2}, \rho, \xi, \beta, \eta, \nu) \right|,$$

i indexes deciles of the marginal distributions, and j indexes the 10 income nodes.

Figure 5.5: The Objective Function Used to Solve for ρ

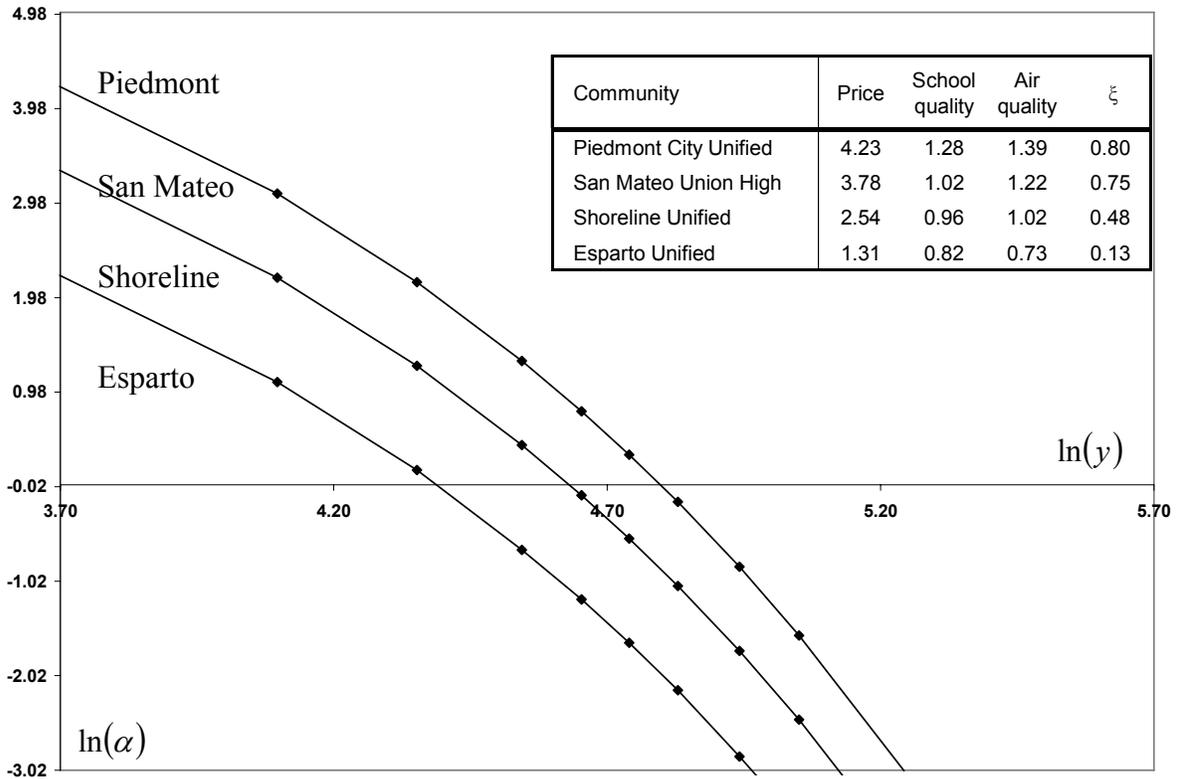


Figure 5.6A: Stratification of Households into Ordered Communities, Horizontal Model

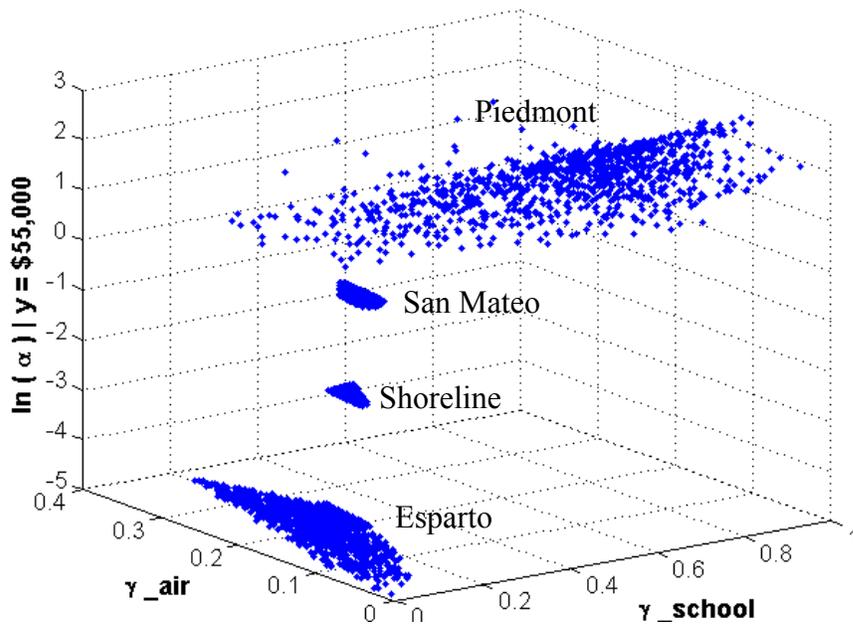


Figure 5.6B: Preference Regions for Households with Income = \$55,000

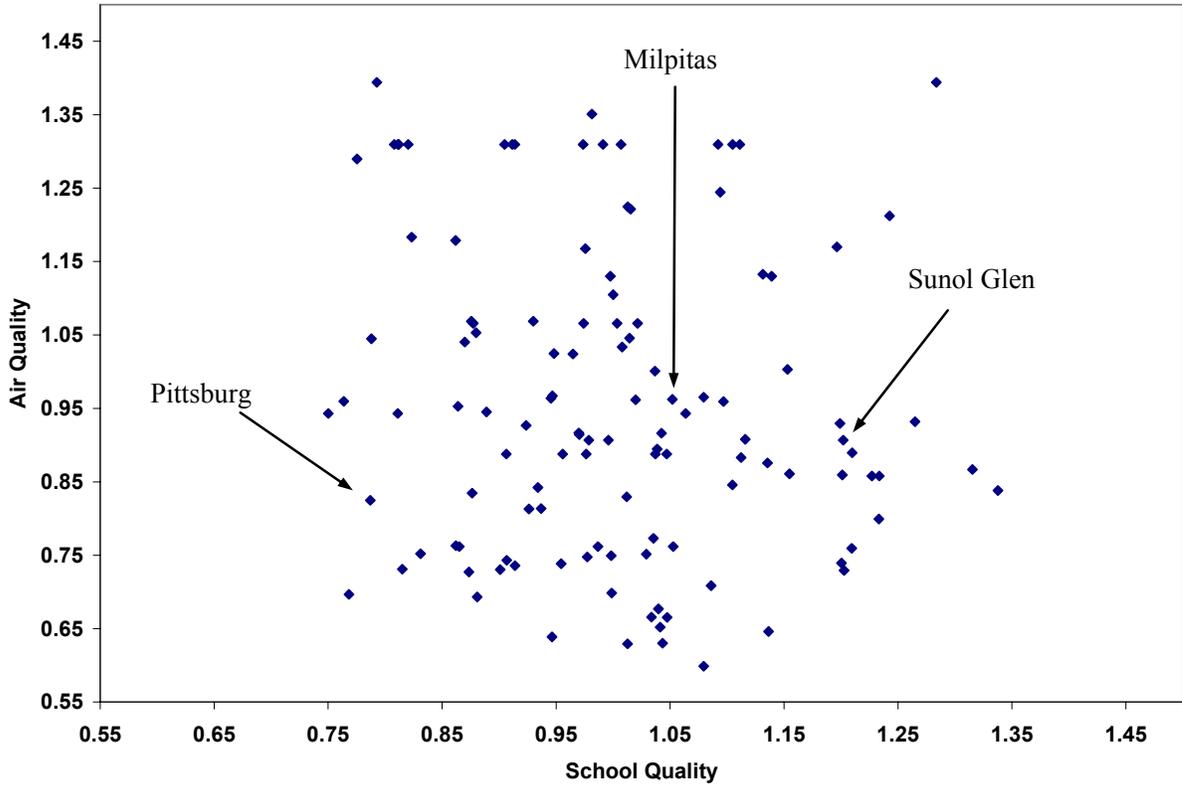


Figure 5.7A: Communities in Attribute Space

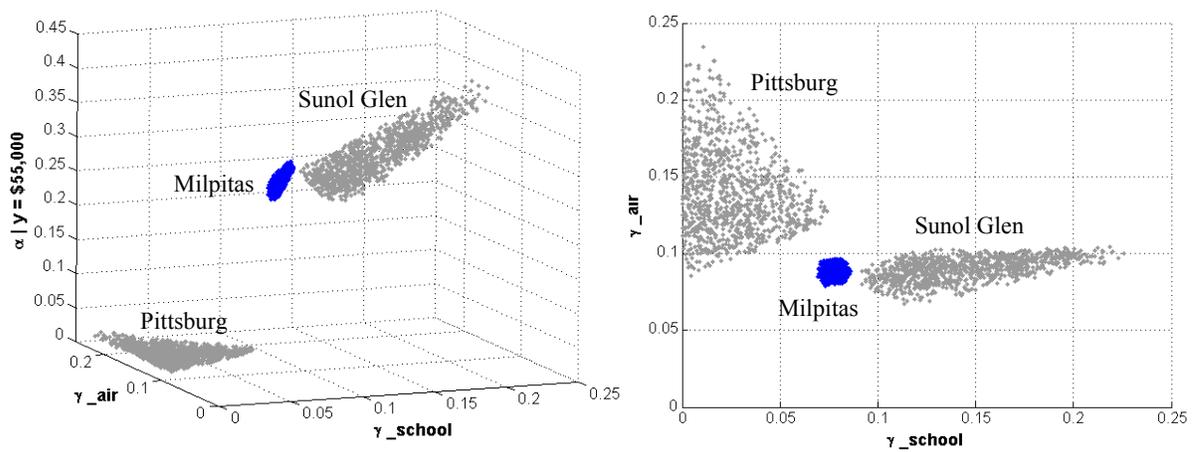
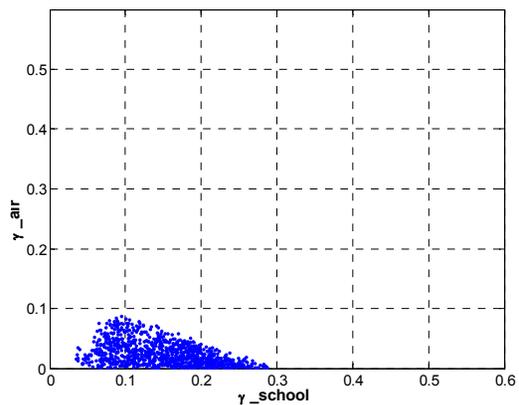


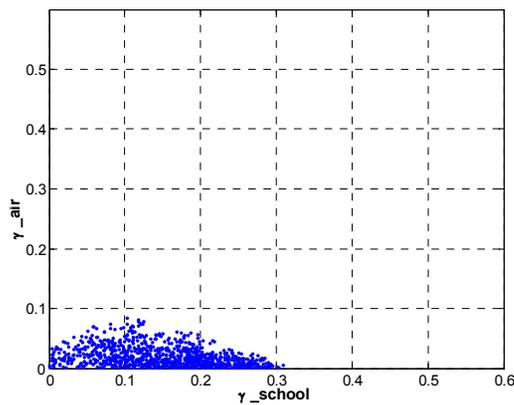
Figure 5.7B: Preferences Regions for 3 Communities, Single-Market Horizontal Model

Architects & Engineers Living in the Acalanes School District

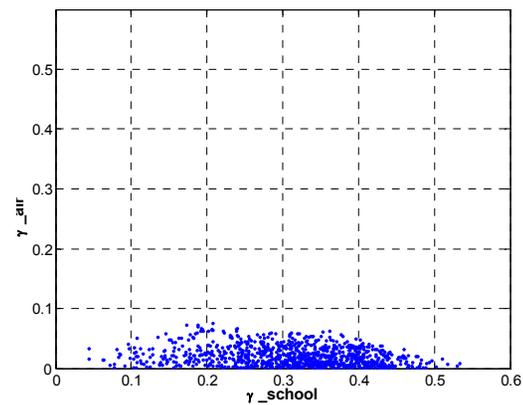
A. Acalanes (single-market)



B. Acalanes → Oakland (24 mins, \$57,700)

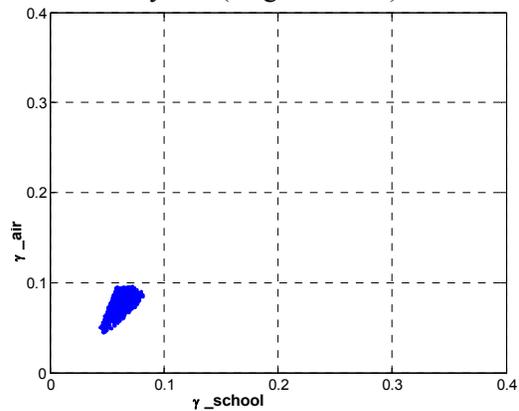


C. Acalanes → San Francisco (55 mins, \$58,500)

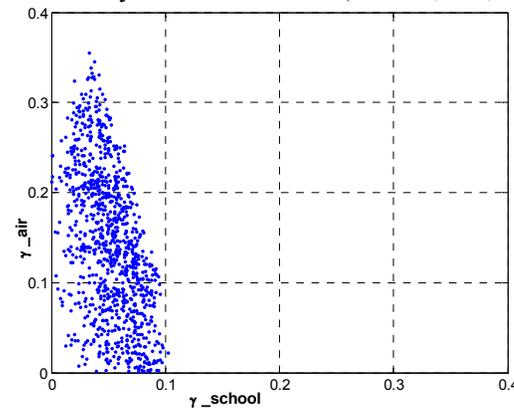


Managers Living in the Berryessa School District

D. Berryessa (single-market)



E. Berryessa → San Jose (23 mins, \$91,100)



F. Berryessa → Oakland (29 mins, \$78,200)

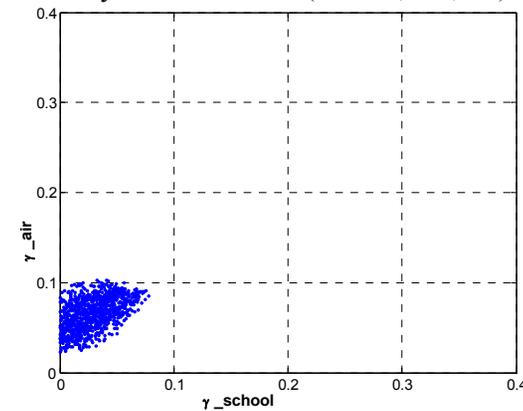
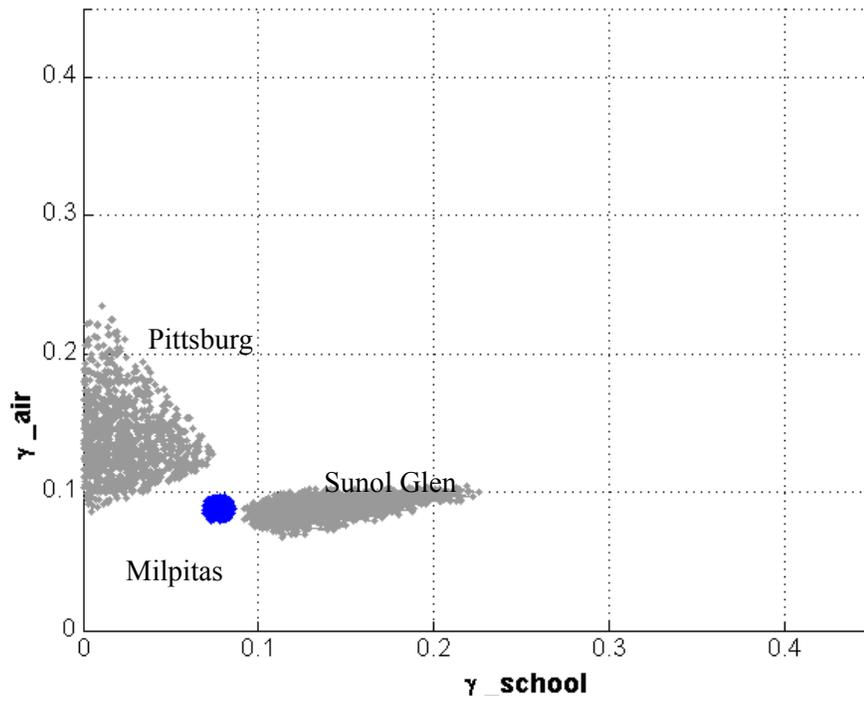


Figure 5.8: Relative Preferences for Air and School Quality in the Single and Dual-Market Versions of the Model

A. Single-Market Partition



B. Dual-Market Partition, Managers

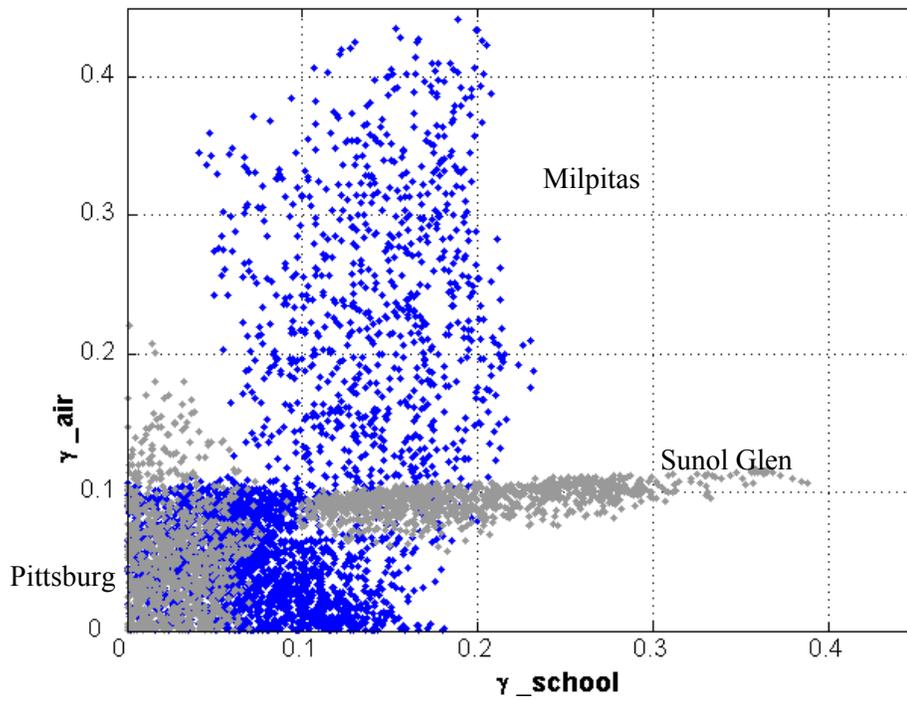


Figure 5.9: Stratification by Relative Preferences with and without Job Opportunities

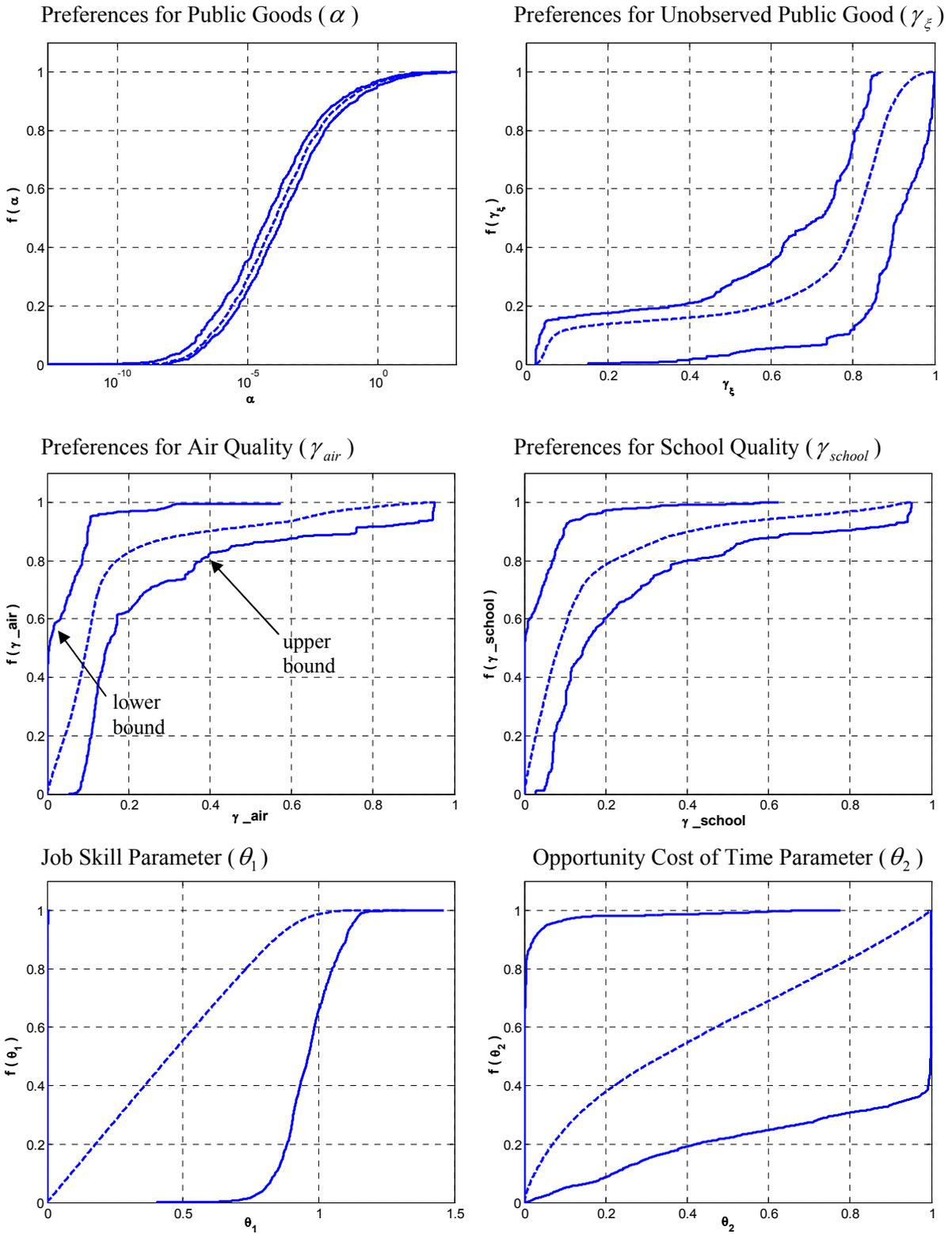


Figure 5.10: Bounds on the Marginal Distribution of Each Heterogeneous Preference Parameter

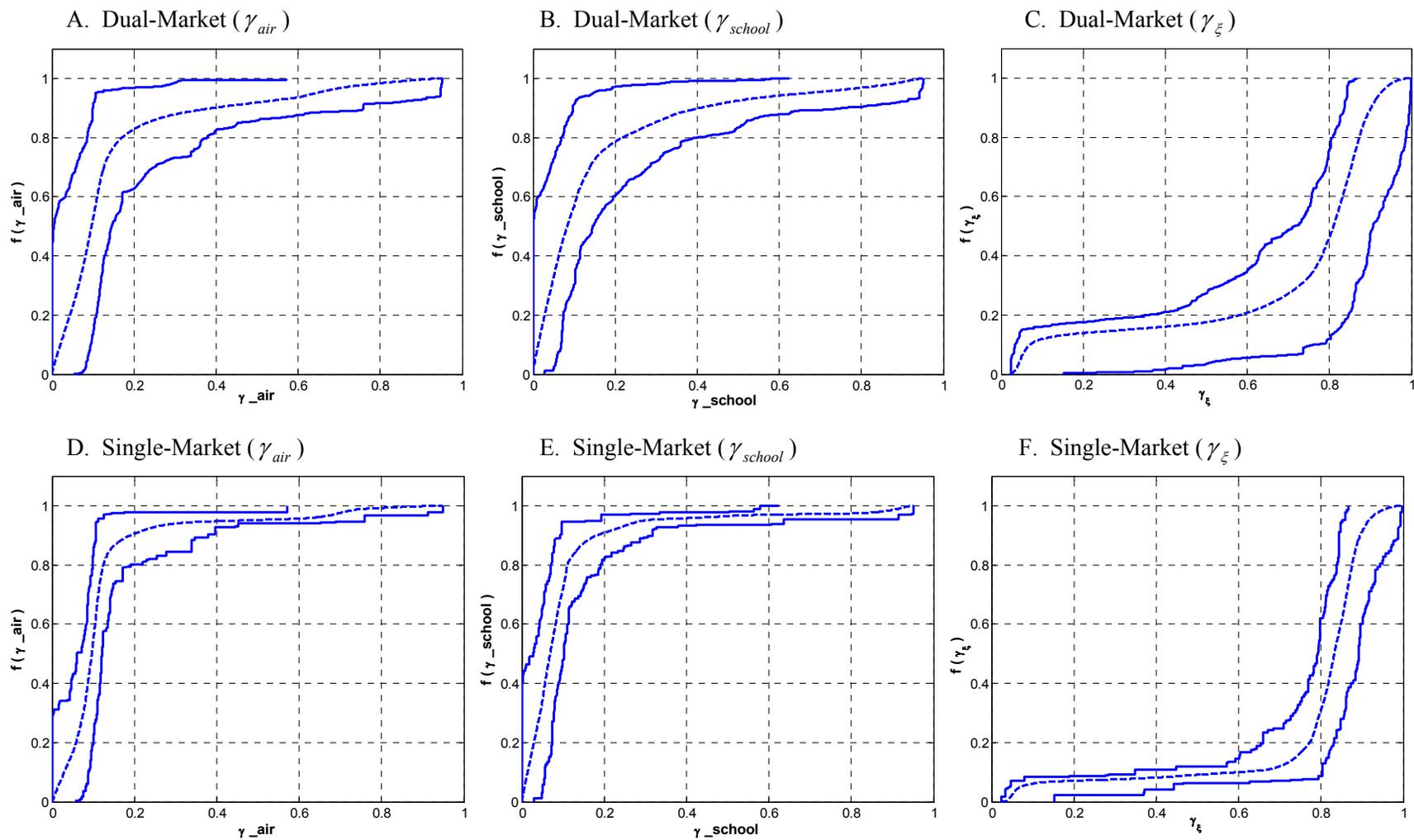
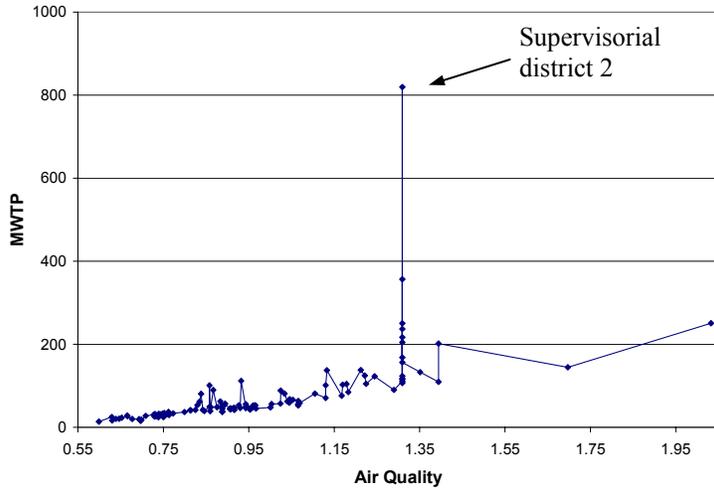
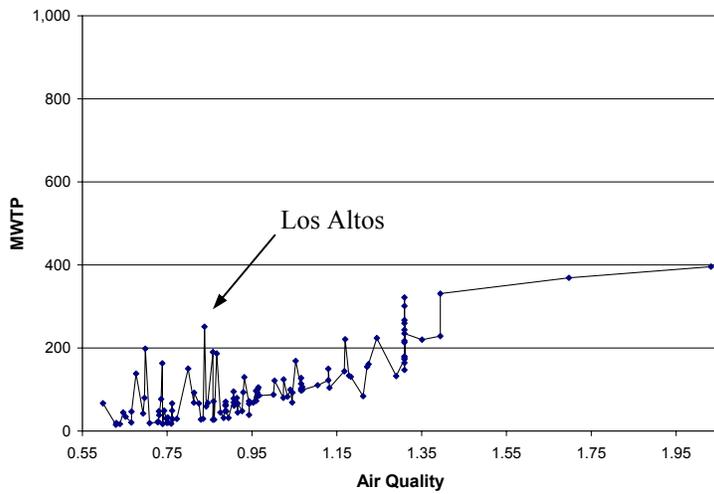


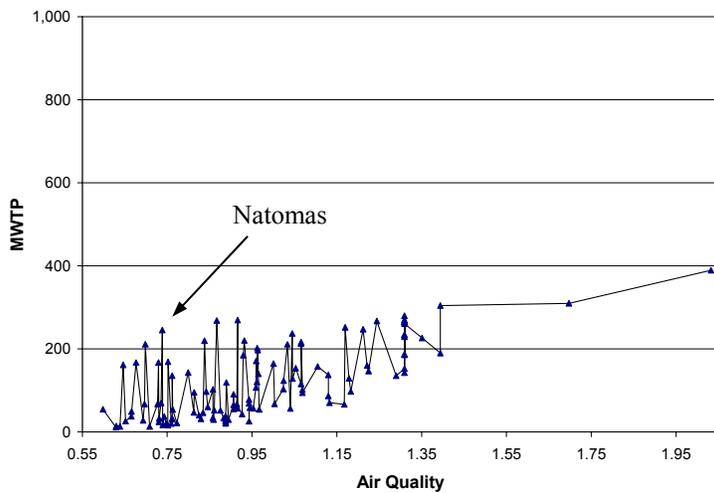
Figure 5.11: Comparison Between Preference Distributions in the Single and Dual-Market Versions of the Model



(1) Single-Market, Vertical



(2) Single-Market, Horizontal
(uniform assumption)



(3) Dual-Market, Horizontal
(uniform assumption)

Figure 5.12: Air Quality and Average MWTP per Community

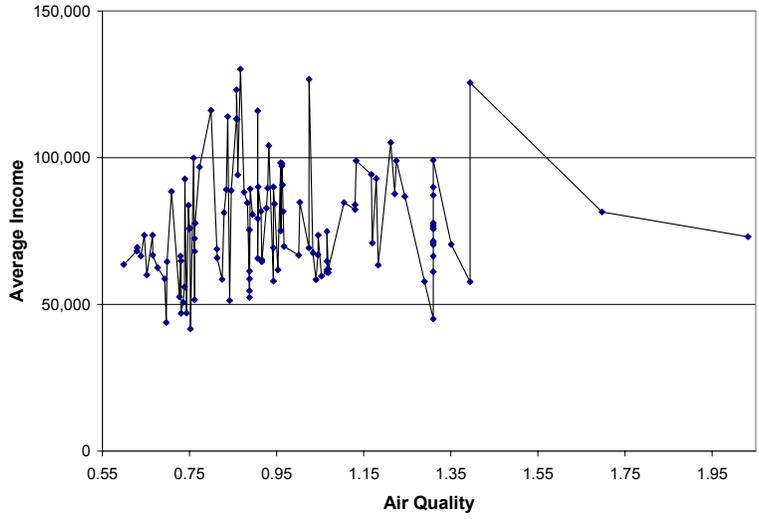


Figure 5.13: Air Quality and Average Income, by Community

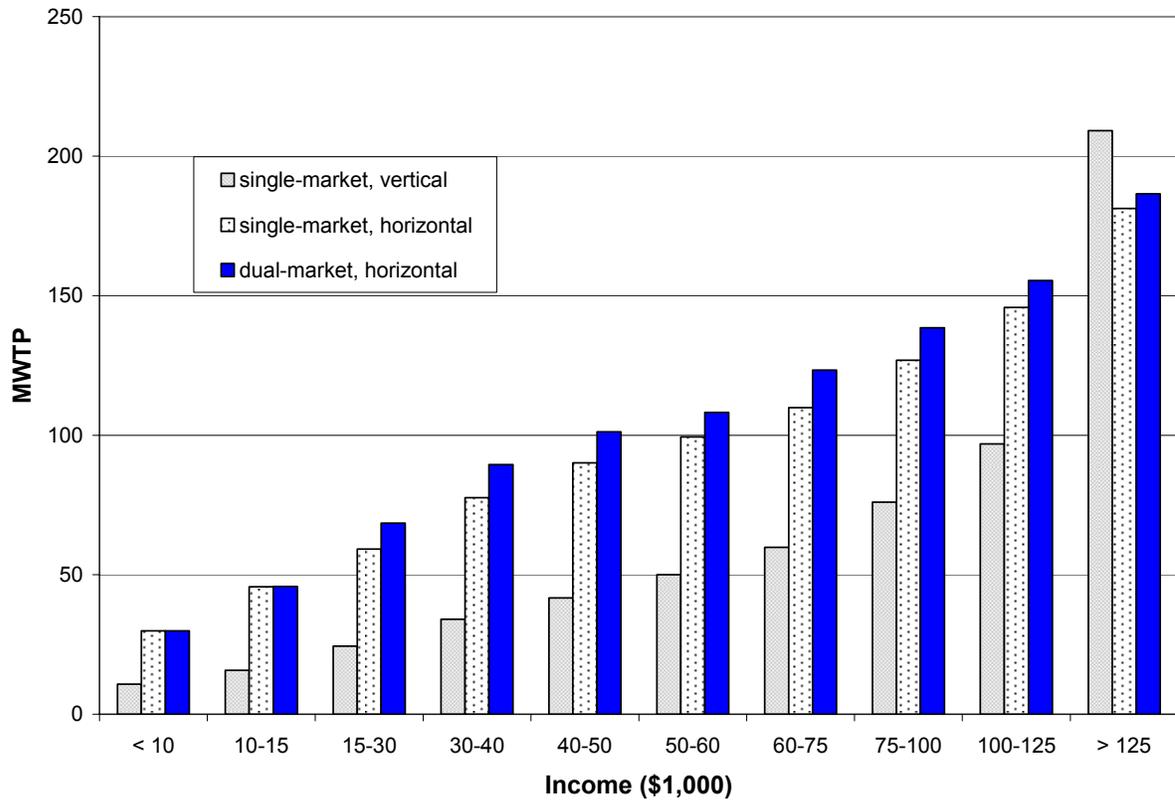


Figure 5.14: MWTP for an Air Quality Improvement by Income

Chapter 6: General Equilibrium Policy Analysis

I. Introduction

This chapter develops a framework to simulate how households and markets adjust to changes in the provision of local public goods and uses it to analyze the welfare implications of California's recently revised standards for ambient concentrations of ground level ozone. The estimation results from the previous chapter can be used to generate a distribution that describes households' *ex ante* willingness-to-pay for an improvement to air quality or school quality in the San Francisco-Sacramento region. However, a sufficiently large improvement to these public goods will induce some households to move to a new house and perhaps to a new job, placing upward pressure on housing prices and downward pressure on wage rates in the improved locations. Subsequent adjustments to prices and wages needed to clear those markets will have additional welfare implications. The first objective of this chapter is to use the information about preferences and job skill recovered during the estimation to simulate the "general equilibrium" adjustment process. The prices, wages, and location choices that define the new equilibrium can then be used to generate a distribution that describes households' *ex post* willingness-to-pay for the improvement.

Recall from chapter 3 that the empirical model generalizes the Epple-Sieg framework to allow wage income to vary with location, and to allow households to differ in their relative preferences for multiple public goods. Likewise, chapter 2 explained how the underlying depiction of the urban landscape generalizes the interregional hedonic literature to recognize that working households may commute between metropolitan areas. These generalizations expand the scope for spatial substitution. In other words, they expand a household's opportunities to adjust its location following a change in public goods. The second objective of this chapter is to investigate the economic consequences of these generalizations. This is addressed by comparing results from four versions of the policy simulation that incrementally expand the scope for spatial substitution, starting with the Epple-Sieg model. The second version generalizes the Epple-Sieg model to allow households to have horizontally differentiated preferences for public goods, and the third version allows workers to change their job locations but requires them to live and work in the same metropolitan area. This commuting restriction is relaxed in the fourth version of the model which solves

for a new set of housing prices, wage rates, and location choices that clear all markets following a change in public goods; i.e. a new “dual-market” locational equilibrium.

Section II starts by formalizing the conditions for a dual-market locational equilibrium and demonstrating how it nests the alternative equilibrium concepts that underlie the other three versions of the model. Sieg, Smith, Banzhaf, and Walsh (2004) previously developed a framework to simulate general equilibrium adjustment in the Epple-Sieg model. However, their approach does not allow horizontally differentiated preferences or labor market adjustment. Section III develops a new simulation framework that incorporates these generalizations and section IV analyzes the properties of the new framework. Finally, section V uses the framework to simulate how households and markets would adjust to a reduction in ozone concentrations that allows the San Francisco-Sacramento region to meet California’s new 8-hour standard. The welfare implications of this adjustment are compared across the four versions of the model.

II. General Equilibrium Concepts and Welfare Measures

Compared to the alternative equilibrium concepts that motivate the hedonic and sorting literatures, the notion of a dual-market locational equilibrium recognizes additional opportunities for spatial substitution. Part (A) of this section starts by defining a dual-market locational equilibrium. The equilibrium conditions are then used to provide definitions for the “partial equilibrium” and “general equilibrium” willingness-to-pay for a change in public goods. Part (B) discusses the four versions of the model used in the policy simulation and explains how the equilibrium concept that underlies each is nested by the dual-market locational equilibrium.

(A) Dual-Market Locational Equilibrium

A dual-market locational equilibrium is characterized by housing prices, wage rates, and location choices such that all markets clear and no household could improve its utility by moving to a new location. Equation (6.1) formalizes these conditions.

$$(6.1.A) \quad H_j^S(p_j) = H_j^D(p_j, w_{1,1}, \dots, w_{s,k_j}) \quad \forall j, \text{ where } k_j \text{ indexes feasible commutes.}$$

$$(6.1.B) \quad L_{s,k}^S(w_{s,k}, p_1, \dots, p_{j_k}) = L_{s,k}^D(w_{s,k}) \quad \forall s, k, \text{ where } j_k \text{ indexes feasible commutes.}$$

$$(6.1.C) \quad V_i(g_j, p_j, w_{s,k}, t_{j,k} \mid \hat{y}_i, \alpha_i, \gamma_i, \theta_i) \geq V_i(g_l, p_l, w_{s,m}, t_{l,m} \mid \hat{y}_i, \alpha_i, \gamma_i, \theta_i) \quad \forall i, l, m.$$

The first equation shows the market clearing condition for housing in community j . The aggregate supply of housing equals the aggregate demand. While supply and demand both depend on the price of housing, the demand is also an increasing function of the wages paid to the workers who live within commuting distance. Equation (6.1B) gives a parallel condition for occupation- s workers in labor market k . The total number of workers who commute to labor market k depends positively on their wage rate and negatively on the price of housing in the communities nearby. Equation (6.1C) depicts the utility maximization condition for a household that chooses to live in community j and labor market k . All households are required to choose their utility-maximizing locations given their non-wage income, preferences, and job skills.

The dual-market locational equilibrium in (6.1) combines the features of the alternative equilibrium concepts that underlie the existing sorting and interregional hedonic literatures. As in the sorting literature, households choose from a finite set of differentiated house locations according to their heterogeneous preferences for the attributes of those locations (e.g. Epple and Sieg [1999]). As in the interregional hedonic literature, a worker may be paid differently if they move between metropolitan areas (e.g. Roback [1982]). In addition to integrating these features, the dual-market concept generalizes both literatures to depict a wider range of spatial substitution possibilities. More precisely, it recognizes that households may choose to live in one labor market and work in another. Chapter 2 explained how this generalization provides an additional source of information to help identify preferences, and chapter 5 illustrated that households who make the choice to commute between labor markets often reveal features of their preferences by the nature of that choice. In a general equilibrium setting, the ability to commute between labor markets increases a

household's range of potential responses to a shock. They can move to a different community, as in the sorting literature. They can change their job and their house simultaneously by moving to a different metropolitan area, as in the interregional hedonic literature. Or, they can choose to remain in the same community (labor market) and commute to a different labor market (community).

Suppose there is an exogenous shock to the provision of public goods. This may induce some households to move to a different community and perhaps to a different job. As the demand for housing and the supply of labor adjust, rents and wages will need to adjust to clear those markets. The welfare consequences of this general equilibrium adjustment can be analyzed by comparing measures for the partial and general equilibrium welfare effects. Equations (6.2) and (6.3) illustrate how the two measures are calculated for a household that responds to the shock by moving from location (j,k) to location (q,r) . The 0 (1) superscripts represent the initial (new) levels of public goods, prices, and wages in the pre (post) shock equilibrium.

$$(6.2) \quad V_{i,j} [g_j^1, p_j^0; \alpha_i, \gamma_i, y_i^0(\theta_i, w_{s,k}^0, t_{j,k}) - WTP_i^{PE}] = V_{i,j} [g_j^0, p_j^0; \alpha_i, \gamma_i, y_i^0(\theta_i, w_{s,k}^0, t_{j,k})].$$

$$(6.3) \quad V_{i,q} [g_q^1, p_q^1; \alpha_i, \gamma_i, y_i^1(\theta_i, w_{s,r}^1, t_{q,r}) - WTP_i^{GE}] = V_{i,j} [g_j^0, p_j^0; \alpha_i, \gamma_i, y_i^0(\theta_i, w_{s,k}^0, t_{j,k})].$$

The (ex ante) partial equilibrium measure (WTP^{PE}) expresses a household's willingness-to-pay for the change in public goods at their initial location, given their initial wage rate and the initial price of housing in their home community. The (ex post) general equilibrium measure (WTP^{GE}) uses the same reference point before the shock, but includes the welfare effect from moving to a new location evaluated at the new price of housing and the new wage rate at that location. In other words, the general equilibrium measure accounts for capitalization of the quality change into housing prices and wage rates.

(B) *Versions of the Dual-Market Locational Equilibrium Used for Policy Analysis*

The model developed by Sieg, Smith, Banzhaf, and Walsh (2004) forms the basis for most

empirical applications that use the conditions for a locational equilibrium to simulate general equilibrium adjustment⁷⁵. They developed an algorithm to solve for a new equilibrium in the housing market, following a shock to the provision of public goods. Smith et al. (2004) use the model to consider the welfare consequences of reduced ozone levels in the Los Angeles air basin, and Walsh (2003) adapts their approach to analyze open space policies in Wake County, North Carolina. Both applications find substantial differences between partial and general equilibrium welfare measures. Unlike the current model, these applications treat income as exogenous and preferences as vertically differentiated. To investigate the consequences of these restrictions in a general equilibrium setting, the policy simulation in section V is conducted for four versions of the model—the dual-market framework in (6.1) and three special cases.

The first two special cases treat income as exogenous and solve for a new equilibrium in the housing market. This reduces the system in (6.1) to the housing and utility maximization conditions in (6.4).

$$(6.4.A) \quad H_j^S(p_j) = H_j^D(p_j) \quad \forall j.$$

$$(6.4.B) \quad V_i(g_j, p_j | y_i, \alpha_i, \gamma_i) \geq V_i(g_q, p_q | y_i, \alpha_i, \gamma_i) \quad \forall i, q.$$

One version replicates the Epple-Sieg framework by treating preferences as vertically differentiated, providing a link to the existing literature. In this case, the population of households can be defined by the joint distribution of income and preferences estimated from the single-market vertical model in chapter 5. The second single-market case allows households to have horizontally differentiated preferences and uses the corresponding partition of preference space from chapter 5 to characterize households.

The partition of preference space that was recovered from the dual-market version of the estimator is used to define the population of households in the other two versions of the policy simulation. Following a shock to the provision of public goods, both models solve for

⁷⁵Bayer et al. (2005) also use their (logit) sorting framework to solve for a new equilibrium. However, chapter 2 explained that the resulting welfare measures do not appear to be theoretically consistent.

a new vector of housing prices and a new matrix of wage rates that simultaneously clear the housing and labor markets⁷⁶. They differ in how they depict the possibilities for spatial substitution. The “free mobility” case defines the choice set using the same 268 (community, PMSA) locations as in the estimation. A second “limited mobility” case restricts households to live and work in the same PMSA. This reduces the choice set to 122 (community-PMSA) combinations. Comparing the results from the two simulations reveals the implications of allowing commuting between labor markets. At a conceptual level, the “limited mobility” case is similar to Roback (1982) and Blomquist, Berger, and Hoehn (1988) in that it requires workers to change their house location if they move to a new job location. However, unlike Roback’s empirical application, workers are free to move between communities within a PMSA; i.e. they can change their house without changing their job. Blomquist et al. incidentally allow workers to move within a PMSA through their use of data⁷⁷.

Generalizing the Epple-Sieg model to allow horizontally differentiated preferences and labor market choices complicates the task of solving for a new equilibrium following a shock to the provision public goods. The next section develops a simulation framework that accommodates these two generalizations.

III. Developing a Framework for General Equilibrium Simulation

In the limit, all the attributes of a location may be endogenous and interrelated in a general equilibrium. For example, by inducing households to move, an air quality improvement could change the composition of the distribution of preferences for school quality within a community. This, in turn, could lead to a change in fiscal allocations to local public schools, affecting the community’s school quality. Likewise, if households respond to the air quality improvement by adjusting their commuting patterns, the corresponding change in vehicle miles traveled could alter the level of automobile emissions, which would feed back into the

⁷⁶ The rows and columns of the wage matrix are occupations and PMSAs.

⁷⁷ Blomquist et al. adopt Roback’s model, treating individual counties as unique (job, house) combinations. However, their wage data are defined at the level of a PMSA, which are comprised of one or more counties. This implies a worker may change their house without changing their job.

level of air quality. To keep the simulation framework manageable, two maintained assumptions are introduced that limit the scope for such interactions. First, provision of public goods is assumed to be exogenous. Therefore, changes in the distribution of income or preferences within a community will not affect its provision of air quality, school quality, or $\bar{\xi}$. Second, there are assumed to be no complementarities in production between workers with different job skills. Therefore, the demand for workers in each occupational category will be independent within a PMSA. In addition to these two (new) assumptions, households are still assumed to be freely mobile.

Given an exogenous shock to the provision of public goods, the general equilibrium “problem” is to solve for a new set of housing prices, wage rates, and location choices that satisfy the conditions for a locational equilibrium in (6.1). This can be done by using the distribution of preferences recovered during the estimation to simulate how households would respond to the shock. However, the estimation results are not sufficient to characterize all the components of the general equilibrium model; some additional information must be provided. Part (A) of this section explains how each component of the model can be defined. Then part (B) explains the solution method that Sieg et al. (2004) developed for the special case where wages are exogenous and preferences are vertically differentiated. When vertical differentiation is relaxed, their approach cannot be used. Part (C) develops an alternative algorithm for the case where households have horizontally differentiated preferences, and part (D) extends the algorithm to the dual-market version of the model. Finally, part (E) concludes by considering issues associated with implementing the new algorithm.

(A) Defining the Components of the General Equilibrium Framework

The general equilibrium model has four sets of components: the supply and demand curves for housing in every community and the supply and demand curves for labor in each PMSA. The housing demand and labor supply curves are implicitly defined by the joint distribution of income and preferences that was estimated for the population of households in the study region in chapter 5. Recall that an individual household’s demand for housing is defined as

follows:

$$(6.5) \quad h_{i,j} = \beta p_j^\eta y_i^\nu .$$

Aggregating (6.5) over all the households in community j gives the quantity of housing demanded in that community. Likewise, the supply of labor for a given occupation is defined by the number of workers in each labor market.

Additional information must be provided to define the supply of housing and the demand for labor. In previous applications, the housing supply curves have been treated as perfectly inelastic (Smith et al. [2004]), calibrated using a range of elasticities (Sieg et al. [2004]), and estimated independently (Walsh [2003]). While any of these approaches could be used here, the policy analysis in section V assumes that the housing supply and labor demand are both perfectly inelastic. Under this assumption, the supply of housing in community j is defined by summing (6.5) over all the households observed living in that community in the (year 2000) census data. Likewise, the demand for labor in each PMSA is defined by the observed working population.

Finally, an assumption is required about who collects the capital gains (or losses) from changes in housing prices and wage rates. The applications cited above all treat households as renters (as opposed to homeowners) and assume that changes in property values are absorbed by absentee landlords. The same assumption is maintained here and extended to the labor market. In other words, changes in profits that arise from adjustments to wage rates are assumed to be collected by the absentee owners of firms (i.e. “shareholders”).

The decision to treat households as renters rather than homeowner/shareholders will affect the welfare implications that emerge from a policy analysis. For example, suppose an improvement in the quality of a public good raises housing prices. This would be strictly welfare improving for a homeowner. The quality improvement increases their utility and the increase in housing prices can not make them worse off. Since homeowners effectively rent from themselves, an increase in housing expenditures is exactly offset by an increase in rental income. Likewise, if the homeowner decides to move, they collect the capital gains

from the increase in the price of their home. In contrast, a renter living in the same community will experience a welfare loss if the increase in their housing expenditures exceeds their willingness-to-pay for the quality improvement. The opposite pattern of results would be produced by a decrease in the price of housing that stems from a decrease in the quality of local public goods. In this case, a homeowner's welfare would decrease whereas a renter would experience a welfare gain if their housing expenditures decrease by more than enough to compensate them for the decreased quality of local public goods. Section IV considers additional implications of the decision to treat households as renters in the context of a specific example.

(B) *Solving for a New Equilibrium: Single-Market Vertical Model*

In the single-market case, the general equilibrium problem requires solving for households' location choices and a vector of prices that clears the housing market in every community following a shock to public goods. When households have vertically differentiated preferences for public goods, the simulation process is relatively straightforward. Recall from chapter 2 that under vertical differentiation, the *ascending bundles* property requires everyone to agree on a common ranking of communities by the public goods they provide, and this ranking must follow the ranking by price. This property allows the problem to be transformed into a one-dimensional rootfinding problem.

Equation (6.6) applies the *boundary indifference* condition to the indirect utility function from the Epple-Sieg model (see equation [2.34])⁷⁸. It defines all the combinations of income and preferences that make a household exactly indifferent between communities j and $j+1$.

$$(6.6) \quad \ln(\alpha_i) - \rho \left(\frac{y_i^{1-\nu} - 1}{1-\nu} \right) = \ln \left(\frac{Q_{j+1} - Q_j}{\bar{g}_j^\rho - \bar{g}_{j+1}^\rho} \right) = R_{j,j+1} \quad ,$$

⁷⁸ Alternatively, this corresponds to the indirect utility function from the empirical model shown in equation (3.1) in the case where $\theta_1 = 1, \theta_2 = 0$, and $\gamma_i = \gamma \forall i$.

$$\text{where } Q_j = \exp\left[-\frac{\rho}{1+\eta}(\beta p_j^{\eta+1} - 1)\right].$$

In the vertically differentiated case, all the heterogeneity in income and preferences appears on the left hand side of the equation; i.e. the values for the index of public goods (\bar{g}) are constant. A household with income and preference such that:

$\ln(\alpha_i) - \rho[(y_i^{1-\nu} - 1)/1 - \nu] < R_{j,j+1}$ will prefer community j to community $j+1$. The single-crossing condition implies that, in equilibrium, that household must also prefer community j to $j+2, j+3, \dots, J$. Therefore the left hand side of (6.6) can be used to systematically sort households across communities. This ordering provides the basis for an algorithm that simultaneously solves for the new equilibrium prices and location choices.

Following the shock to public goods, the ascending bundles property requires the new price ranking of communities to be the same as the new ranking by the public goods index. Given a guess for the new price of housing in the cheapest community (\tilde{p}_1), the left hand side of (6.6) can be used to sort households into community 1 until the housing market clears. The preferences and income of the final “border” household are then used to solve (6.6) for the price of housing in community 2. The resulting value for p_2 is used to sort the next group of households into community 2 which generates a value for p_3 , and so on. Equation (6.7) depicts the structure of the problem, using star superscripts to indicate market clearing.

$$(6.7) \quad p_2^* = f(\tilde{p}_1) \rightarrow p_3^* = f(p_2^*) \rightarrow \dots \rightarrow p_J^* = f(p_{J-1}^*) .$$

This recursive structure effectively reduces the simulation to a one-dimensional problem where the new equilibrium price of housing in community 1 is adjusted until the market clears in community J . Excess supply in community J indicates that p_1 has been set too high and excess demand indicates it has been set too low.

Unfortunately, the general equilibrium problem cannot be defined recursively if households differ in their relative preferences for multiple public goods. Horizontal

differentiation introduces two interrelated complications. First, the new equilibrium price ranking cannot be predetermined from the new allocation of public goods because households differ in how they rank communities by public goods. Second, there is no longer a unique one-dimensional ordering of all households across all communities. With multiple dimensions of preference heterogeneity there will be multiple dimensions of boundary indifference in the partition of preference space. For example, suppose that community 2 shares preference “borders” with communities 1, 3, and J. This creates circularity in the structure of the problem, as shown in equation (6.8).

$$(6.8) \quad \boxed{\rightarrow p_2^* = f(p_1, p_3, p_J) \rightarrow p_3^* = f(p_2) \rightarrow \dots \rightarrow p_J^* = f(p_{J-1}) \leftarrow}$$

A marginal increase (decrease) in p_J will induce border households to move into (out of) community 2, initiating a new round of price adjustments. This problem cannot be solved recursively without knowing the new price ranking of communities.

Fortunately, even without the recursive shortcut, it is still possible to solve for a new equilibrium by utilizing the necessary conditions for market clearing and utility-maximization. However, the sorting process can become computationally intensive. The mechanics of the new sorting algorithm are discussed before extending it to the labor market.

(C) *Solving for a New Equilibrium: Single-Market Horizontal Model*

When households have horizontally differentiated preferences, the market clearing price of housing in each community generally depends on the price of housing in every other community. Equation (6.9) illustrates this J -dimensional problem.

$$(6.9) \quad p_1^* = f(p_2, \dots, p_J), \quad p_2^* = f(p_1, p_3, \dots, p_J), \dots, \quad p_J^* = f(p_1, \dots, p_{J-1}).$$

While there is no obvious way to sort households across communities, they can still be ranked according to the maximum price they would be willing to pay to live in every community other than their current location. Equation (6.10) defines this *Hicksian Equivalent Price* (HEP).

$$(6.10) \quad p_{i,j}^{HEP} : V_{i,j}(g_j, p_j^{HEP}; \alpha_i, \gamma_i, y_i) = V_{i,q \neq j}(g_q, p_q; \alpha_i, \gamma_i, y_i).$$

Suppose household i lives in community q . Then $p_{i,j}^{HEP}$ is the price of housing in community j that would allow household i to maintain its current level of utility, were it to move there. In a locational equilibrium, the price of housing in j must be greater or equal to the maximum HEP for any household living outside that community: $p_j^* \geq \max \{ p_{i,j}^{HEP} \}$. This necessary condition provides the basis for the new sorting algorithm.

Following the shock to public goods, the algorithm starts by setting the new price of housing in each community equal to the maximum HEP. For example, suppose that the quality of public goods has improved in community 1. Outside residents will be willing to pay more to live there, in which case p_1 may need to increase to keep the housing market in equilibrium. The algorithm solves (6.10) for every household living outside community 1 and then sets the price in community 1 equal to the maximum HEP. Since the demand for housing is downward sloping, an increase in p_1 will decrease the quantity of housing consumed by current residents, creating room for the household with the highest HEP to move in. If the supply of housing still exceeds the demand, the algorithm moves to the household with the next highest HEP and this process continues until the market clears. The new market-clearing price (p_1^1) is then used to update the utility of every household in community 1. (Since all households are treated as renters they all have to pay the increased rental rate of housing.) After updating the utility of residents in community 1, the algorithm repeats the process to update housing prices in communities 2 through J , and the utility of the households who live there.

The first round of adjustments is unlikely to result in a new locational equilibrium. Some of the households who were initially sorted into community 1 may have been subsequently sorted into other communities, leaving community 1 with an excess supply of housing. In this case p_1 must decrease, initiating a second round of adjustments. The algorithm continues to iterate over subsequent rounds of adjustment until all prices converge, signaling that markets have cleared and that each household occupies its utility maximizing location. Equation (6.11) summarizes the steps required to find the new locational equilibrium, using superscripts to represent the current round of adjustments.

(6.11) The HEP Sorting Algorithm

$$\begin{aligned}
 \text{(i)} \quad p_1^1 &= \max \{ p_{i,1}^{HEP} \} = f(p_2^0, \dots, p_J^0). \\
 \text{(ii)} \quad p_2^1 &= \max \{ p_{i,2}^{HEP} \} = f(p_1^1, p_3^0, \dots, p_J^0). \\
 &\vdots \\
 &\vdots \\
 \text{(J)} \quad p_J^1 &= \max \{ p_{i,J}^{HEP} \} = f(p_1^1, \dots, p_{J-1}^1). \\
 \text{(J+i)} \quad p_1^2 &= \max \{ p_{i,1}^{HEP} \} = f(p_2^1, \dots, p_J^1). \\
 \text{(J+ii)} \quad p_2^2 &= \max \{ p_{i,2}^{HEP} \} = f(p_1^2, p_3^1, \dots, p_J^1). \\
 &\vdots \\
 &\vdots \\
 \text{(2J)} \quad p_J^2 &= \max \{ p_{i,J}^{HEP} \} = f(p_1^2, \dots, p_J^2). \\
 &\vdots \\
 &\vdots \\
 &\vdots \\
 &\text{until all prices converge.}
 \end{aligned}$$

After solving for the new equilibrium, the welfare consequences of general equilibrium adjustment can be analyzed by comparing measures for the partial and general equilibrium welfare effects.

(D) Solving for a New Equilibrium: Dual-Market Horizontal Model

It is relatively straightforward to extend the HEP algorithm to allow households to adjust their job locations and to allow wages to adjust. The same logic used to define the Hicksian equivalent price can be used to define a *Hicksian Equivalent Wage* (HEW) that would make a working household exactly indifferent between its current location and another location in a different labor market.

$$(6.12) \quad w_{s,k}^{HEW} : V_{i,k}[\mathbf{g}, \mathbf{p}; \alpha_i, \gamma_i, y_i(\theta_i, w_{s,k}^{HEW}, t_k)] = V_{i,r \neq k}[\mathbf{g}, \mathbf{p}; \alpha_i, \gamma_i, y_i(\theta_i, w_{s,r}, t_r)].$$

In the equation, $w_{s,k}^{HEW}$ is the wage in labor market k that would allow a household with occupation s to change its job location to k and maintain the same level of utility it gets from its current location in labor market r . Community-specific subscripts are suppressed in the equation because working households are assumed to choose the utility-maximizing housing community conditional on their job location. A necessary condition for locational equilibrium is: $w_{s,k}^* \leq \min\{w_{i,s,k}^{HEW}\}$. That is, there cannot be any households in alternative job locations who would be willing to work for less than the current market wage.

Simultaneously solving for a new set of housing prices and wage rates simply requires adding another loop to the HEP sorting algorithm that solves for the HEW for each occupation in each labor market. Equation (6.13) summarizes the steps in the algorithm. It iterates over updates to housing prices given wages and updates to wages given housing prices.

(6.13) The HEP-HEW Sorting Algorithm

$$\begin{aligned} \text{(i)} \quad & p_1^1 = \max\{p_{i,1}^{HEP}\} = f(p_2^0, \dots, p_J^0 | w^0). \\ \text{(ii)} \quad & p_2^1 = \max\{p_{i,2}^{HEP}\} = f(p_1^1, p_3^0, \dots, p_J^0 | w^0). \\ & \vdots \\ & \vdots \\ \text{(J)} \quad & p_J^1 = \max\{p_{i,J}^{HEP}\} = f(p_1^1, \dots, p_{J-1}^1 | w^0). \end{aligned}$$

$$\begin{aligned}
\text{(J+i)} \quad & w_{1,1}^1 = \min\{ w_{1,1}^{HEW} \} = f(w_{1,2}^0, \dots, w_{1,K}^0 | p^1). \\
& \cdot \\
& \cdot \\
& \cdot \\
\text{(J+SK)} \quad & w_{S,K}^1 = \min\{ w_{S,K}^{HEW} \} = f(w_{S,1}^0, \dots, w_{S,K-1}^0 | p^1). \\
\text{(J+SK+i)} \quad & p_1^2 = \max\{ p_{i,1}^{HEP} \} = f(p_2^1, \dots, p_J^1 | w^1). \\
& \cdot \\
& \cdot \\
& \cdot \\
\text{(2J+2SK)} \quad & w_{S,K}^2 = \min\{ w_{S,K}^{HEW} \} = f(w_{S,1}^1, \dots, w_{S,K-1}^1 | p^2). \\
& \cdot \\
& \cdot \\
& \cdot \\
& \text{until all prices and wages converge.}
\end{aligned}$$

Of course, for retired households income is still exogenous to location choice. They only enter the steps of the algorithm which adjust prices. The set of housing prices, wages, and location choices that define the new locational equilibrium can be used to calculate general equilibrium welfare effects that incorporate income changes in addition to capitalization in housing prices.

(E) Implementation Issues

Implementing the new algorithms can be computationally intensive and, in the dual-market case, can produce multiple equilibria. Unlike the recursive approach used in the vertically differentiated scenario, the HEP and HEP-HEW algorithms use “brute force” to solve for a new equilibrium. For example, notice that the HEP algorithm calculates almost every household’s HEP on every step of (6.11). One strategy to reduce the computational burden would be to define the population of households as a fraction of its true size. For example, a 1-in-100 sample would assume that each draw from the joint distribution of income and preferences represents 100 households. However, this would reduce continuity in the joint distribution. A second strategy would be to use data on households from a subset of locations that provide reasonably complete coverage of the study region. Like the sub-

sampling strategy, this would reduce continuity in the distribution of preferences. Moreover, by reducing the possibilities for spatial substitution, it may fail to capture important relationships between locations that are included in the simulation and locations that are excluded.

Allowing labor markets to adjust raises the possibility of multiple equilibria. Intuitively, an increase in wage income can always be offset in the housing market by a corresponding increase in prices. This can lead to degenerate solutions where prices and wages increase indefinitely. The stylized diagrams in figure 6.1 illustrate the problem. Suppose that, following an air quality improvement, the price of housing in community j increases from p_j^0 to p_j^1 . Households who live in that community, work nearby, and are indifferent to air quality may respond to the price increase by moving to a different (job, house) combination in a different metropolitan area. Suppose that the occupation- s workers who moved out of community j formerly worked in labor market k . Now suppose that the households who move into community j have jobs in different occupational categories. Therefore, the wages paid to type- s workers must increase to $w_{s,k}^1$ to ensure the labor market clears. Assuming there are still some type- s workers living in community j , the income effect from the wage increase will lead to excess demand in housing market j , requiring the price of housing to increase to p_j^2 , leading to $w_{s,k}^2$ and so on.

One strategy to stop the cycle of price and wage increases would be to impose a limit on the maximum wage that could be paid to workers in each occupation. This could be interpreted as a zero profit condition for firms or, alternatively, as the “entry wage” that would induce workers to move in from outside the study region. The next section uses an example to test this strategy and to test the two strategies for decreasing the computational burden of the simulations.

IV. Using an Example to Investigate Implementation Issues

Initial attempts to implement the new algorithms indicated that computational time is likely to pose a binding constraint in practical applications. For example, using a 100% sample of the 3.2 million households in the study region, the HEP algorithm was unable to converge within 5 days on a standard desktop computer. Likewise, the dual-market simulation was characterized by a divergent cycle of price and wage increases. This section presents results from investigating practical strategies to address these two issues. Part (A) presents results from running the single-market simulation using a subset of communities and a fraction of the population, and then discusses the consequences of these simplifications for welfare measures. Part (B) presents results from running the dual-market simulation using a subset of locations and a fraction of the population, while imposing a ceiling on the wage that can be paid to workers in each occupation. Finally, part (C) summarizes the results and discusses caveats to using the new algorithms for policy analysis.

(A) Computational Burden

Given the difficulty in running a full scale version of the simulation, the two strategies for reducing the computational burden were each evaluated conditional on the other. The consequences of using a subset of the population were explored by comparing the simulation results from 1-in-100, 1-in-10, and 1-to-1 samples of the population, using a subset of 20 communities. Then, the consequences of using a subset of 20 communities were explored by comparing the results from the 20-community 1-in-100 sample with the results from another 1-in-100 simulation that used all 122 communities. Throughout these exploratory exercises, the population of households used in the simulation was drawn using the results from the “single-market horizontal” distribution of preferences that was created in chapter 5 by sampling uniformly over the partition of preference space.

(i) Using a Subset of the Population

Table 6.1 summarizes the sensitivity of the simulation results to sample size. For each of the 20 communities, the table shows the simulated shock to ozone levels, the subsequent change in housing prices, the share of the initial population that moved during the simulation, and the average MWTP, WTP^{PE} , and WTP^{GE} for the households that started in each community. These statistics are reported for 1-in-100, 1-in-10, and 1-to-1 sample populations. As chapter 5 discussed, the pattern of results for the average MWTP is consistent with the idea that, all else held constant, households with stronger preferences for air quality tend to sort into communities with cleaner air (i.e. lower ozone). Notice that increasing the sample size has very little effect on the MWTP and WTP^{PE} . This simply indicates that for most communities a 1% sample is sufficiently large to obtain a reasonably accurate estimate for the mean values of the preference parameters in most regions of the partition of preference space.

To start the simulation, the initial equilibrium was shocked by reducing ozone concentrations to 76 parts per billion (ppb) in the five school districts with the highest ozone concentrations. The percentage decrease ranges from 3% in the Antioch school district to 28% in the Black Oak Mine school district. Following this shock, the share of households who move ranges from 0% to 10.5% depending on the community and the sample size. The “price change” columns in the table show that increasing the sample size systematically decreases the magnitude of the adjustment in housing prices. This affects the size and direction of the resulting welfare measures in the WTP^{GE} columns. With a 1% sample, the price of housing increases in every community, causing the average household’s welfare to *decrease* in 17 of the 20 communities⁷⁹. Some of these price increases are counterintuitive. For example, housing prices increase in Alum Rock even though air quality is unchanged and nobody has moved in or out. The reason is that some households in Acalanes would prefer to move to Alum Rock following the price increase in Acalanes. These households bid up the price of housing in Alum Rock, but not by enough to induce any of the current residents to leave. In contrast, when the sample size is increased to 10% of the population the only

⁷⁹ Recall that, *ceteris paribus*, an increase in the price of housing leads to a welfare loss because households are treated as renters rather than homeowners.

communities with price increases are those which experienced air quality improvements. Curiously, the price of housing actually decreases in one of the improved communities—Antioch. There are two reasons why this occurs. First, the pattern of ozone reductions makes Antioch relatively less attractive than the other communities with air quality improvements. Second, the large price decrease in Cloverdale induces some households to move there from Antioch, requiring its price to decrease by 0.1% to clear the market.

Three features of the simulation can explain the observed sensitivity of capitalization patterns and welfare measures to the sample size. First, increasing the sample size increases the extent to which the population of households covers the partition of preference space. With a relatively sparse distribution of preferences in the 1% sample, there is greater scope for prices to increase without inducing current residents to move. Graphically, this argument is analogous to the discussion of moving costs in chapter 3. The farther households are located from the borders that delineate their region of preference space, the higher the price they can be charged before they move⁸⁰. Larger population sizes will include more “border” households who will respond to price increases by moving.

The second feature of the simulation that contributes to the observed capitalization patterns is that increasing the sample size improves the accuracy of the (approximate) market clearing condition for housing. With a finite sample and a perfectly inelastic supply, some housing markets do not clear exactly. In these markets the HEP algorithm reaches a point where there is still excess supply, but lowering the price enough to induce the marginal household to move in would lead to excess demand. In this situation, the algorithm errs on the side of excess supply, raising the price of housing above the level that would exactly clear the market. Increasing the sample size also increases the total stock of housing in the simulation, diminishing the extent to which the marginal housing transaction can influence prices. With a 1% sample, the median excess housing supply across the 20 communities is 0.3% of the housing stock. The maximum excess supply, however, is 8%. The maximum drops to 0.7% in the 10% sample and then to 0.3% in the 100% sample.

⁸⁰ In the context of a non-durable differentiated product, this would imply lower cross-price elasticities and greater market power for suppliers.

The effects of excess supply in the housing market and discreteness in the distribution of preferences are both magnified by the fact that all communities are substitutes. A higher price of housing in one community increases the scope for prices to increase in other communities; or alternatively it reduces the pressure for prices to decrease.

While moving from a 1% to a 10% sample appears to have a large effect on the general equilibrium welfare measures, the additional change from moving to a 100% sample is relatively small. The median difference in the average WTP^{GE} across communities between the 1-in-10 and 1-to-1 samples is 10.32%. This suggests that, for the current application, a 1-in-10 sample may be a reasonable compromise between accuracy and computational time.

(ii) Using a Subset of Communities

To explore the consequences of using a subset of communities, the simulation was repeated using all 122 communities with a 1-in-100 sample⁸¹. Moving from 20 to 122 communities increases the sample size from 62,000 households to 320,000. The same pattern of ozone reductions was used for the 20 communities in table 6.1, and a similar pattern was created for the other 102. That is, ozone levels were reduced to 76 ppb in the 25 communities with the lowest air quality (out of the 102). Thus, the two simulations differed in their depiction of the choice set and in the sample of households, but were based on a similar pattern of ozone reductions.

The results reported in table 6.2 suggest that, conditional on the 1-in-100 sampling ratio, using a subset of communities will systematically overstate the new price, understate the share of households who move, and understate the general equilibrium willingness-to-pay for an improvement. These differences reflect the limitations on substitution possibilities imposed by excluding 102 communities.

In general, when the price of housing in a community increases, some of the (renter) households who live there will prefer to move. The corresponding decrease in demand requires housing prices to decrease in order to clear the market. This effect limits the extent

⁸¹ This took approximately 6 hours to converge on a standard desktop computer.

to which prices can increase following an improvement in public goods. When the choice set is reduced to 20 communities there is less scope for households to adjust to higher prices by moving. Without their preferred substitute locations, households will be willing to endure larger price increases before they decide to move. This effect is analogous to a decrease in the absolute magnitude of own-price elasticities that would occur when some of the choice alternatives are removed from a demand system for a differentiated product.

Black Oak Mine provides an example of this effect. Within the 20-community subset, Center is the closest substitute for Black Oak Mine. In other words, using Black Oak Mine as a reference point, Center offers the closest price of housing and the closest levels of every public good. However, out of the remaining 102 communities there are 8 with closer levels of school quality, 7 with closer levels of air quality, and 1 with a closer level of $\bar{\xi}$. In the 122-community simulation, 36% of Black Oak Mine's population moves to one of these communities, compared to 4% of the population that moves to Center. Consequently, the price of housing in Black Oak Mine increases less than in the 20-community example, and the WTP^{GE} is larger.

(B) Multiple Equilibria

The HEP-HEW algorithm was tested using the same 20 communities and the same ozone reductions reported in table 6.1, without restricting workers' mobility. In total, there were 46 (community, PMSA) locations. The population of households was defined using the results from the "dual-market" distribution of preferences that was created in chapter 5 by sampling uniformly over the partition of preference space. In order to test for the possibility of multiple equilibria, two versions of the simulation were run. They differed only in the order in which they adjusted prices and wages; one version started by adjusting prices first and the other started by adjusting wages. Both used a 1-in-4 sample of the population and limited the maximum wage for each occupation to the maximum observed in the initial equilibrium for any PMSA. For example, the maximum wage that could be paid to managers in any PMSA was capped at \$91,100—the wage rate observed for the San Jose PMSA in 1999 (table 4.5).

While capping the maximum wage for each occupation allowed the simulation to

converge, the general equilibrium welfare measures reported in table 6.3A depend on the order of price and wage adjustment. When wages adjust first, the simulation tends to produce larger changes in both wages and prices. The net effect on WTP^{GE} ranges from 0.07% for the average household in Alexander Valley to 48% for the average household in Campbell. While the difference in WTP^{GE} for the average household is relatively small (4.8%), these results illustrate the general problem that there is not necessarily a unique equilibrium in the dual-market version of the simulation.

The biggest differences between the general equilibrium welfare measures in table 6.3 and those from the single-market simulations in tables 6.1 and 6.2 occur in the communities with ozone reductions. While the dual-market simulation produces higher housing prices for those five communities than the single-market simulation, wages also tend to increase in nearby labor markets. The net effect is larger measures for WTP^{GE} . For example, workers can commute to the Sacramento PMSA from 3 of the 20 communities: Black Oak Mine, Center, and Davis. Each of these communities experienced an air quality improvement and an increase in the price of housing. Meanwhile, wages increased by \$3,322 for the average worker in Sacramento (in the version of the simulation where wages adjust first). These partly offsetting changes reflect the pattern of price and wage increases depicted in figure 6.1.

The cycle of reinforcement between price and wage increases can be demonstrated by analyzing the location choices made by workers in managerial occupations who started the simulation living in Black Oak Mine and working in the Sacramento PMSA. Prior to the ozone reduction, Black Oak Mine had the lowest rent of the 20 communities. When the rental rate increases in Black Oak Mine following the ozone reduction, many of the managers would prefer to move elsewhere. Therefore, the wage paid to managers in Sacramento must increase to clear the market. The income effect from the higher wage rate increases the managers' demand for housing in Black Oak Mine. This requires further price increases to clear the market for housing, inducing further wage increases, and so on. As the cycle continues, many of the managers move to other communities in Sacramento, alleviating the pressure on rents in Black Oak Mine. In the new equilibrium, wages are \$9,200 higher and

55% of the managers have moved to other communities, but all continue to work in Sacramento. The increase in wages retains the current group of managers rather than luring others from different PMSAs because managers are generally paid more in other PMSAs. For example, in the new equilibrium, the average managerial wage in the San Francisco PMSA is \$4,400 higher.

Despite a higher average wage in Sacramento, wages decrease for workers in 6 of the 22 occupations. Table 6.3B reports the minimum, maximum, and average change in wages within each PMSA. In addition to having the largest wage increase for any occupation (managers: \$9,200), Sacramento also has the largest wage decrease (Arts, Design and Media: -\$1,024). Job mobility appears to play an important role in determining whether wages increase or decrease for each occupation. All else held constant, workers who are more mobile (i.e. those with higher values for θ_1) have more leverage to command a higher wage. Since they will be paid more in alternative locations, they are more likely to move, requiring wages to increase. Figure 6.2 illustrates this by graphing kernel approximations to two probability density functions for the job skill parameter: one for the workers in occupations that experienced wage increases and one for workers in occupations that experienced wage decreases. The workers with wage increases tend to be more mobile. This is especially true in the right tail of the distribution which contains the workers best suited to move.

Finally, table 6.3C decomposes the values for the average WTP^{GE} in each community into the averages for retired and working households. While the average working household experiences a welfare increase in 15 out of the 20 communities, welfare decreases for the average retired household in all but three communities. Since their income is exogenous, retired households lose from the cycle of rent and wage increases.

(C) Summary and Caveats

The results from the single-market simulations suggest that using a subset of communities and a fraction of the population will lead to a new equilibrium in which prices are overstated and general equilibrium welfare measures are understated relative to using all the communities and a 1-to-1 sample of the population. These results imply that welfare

measures based on the results of the HEP algorithm can be interpreted as consistent underestimates of the willingness-to-pay for an improvement in public goods.

In the dual-market case, welfare measures appear to be driven by a cycle of rent and wage increases that effectively transfer wealth from retired households and absentee shareholders to working households and absentee homeowners⁸². As a result, wages increase for the average worker in every PMSA following the ozone reduction. This contradicts the conventional logic from the interregional hedonic model (Rosen [1979]) which implies that workers would have to “pay” for amenity improvements through lower wages. Figure 6.2 illustrates how the wage increases are tied to high values for the job mobility parameter. In contrast, if all households had sufficiently low values for θ_1 , they would never change jobs and wages would remain the same or decrease. While the wage increases stem from heterogeneity in job skill, the magnitude of those increases is tied to the modeling decision to treat households as renters rather than owner/shareholders. If households collected the capital gains from increases in rents, they would not need such large wage increases to keep them from changing jobs. Likewise, making households “pay” for wage increases through lower non-wage income (i.e. lower dividends from ownership shares in firms) would counterbalance the wage-based income effect on the demand for housing. In an extreme case, the simulation could be run as a zero-sum game where wealth never enters or leaves the model. This is left as a possibility for future research.

To the extent that households are actually homeowners/shareholders the cycle of rent and wage increases will cause the HEP-HEW algorithm to overstate WTP^{GE} for working households and understate WTP^{GE} for retired households. For retired (working) households, this would reinforce (contradict) the effects from using a subset of communities and a fraction of the population in the simulation. Furthermore, because there can be multiple equilibria, welfare measures based on the HEP-HEW algorithm may be sensitive to seemingly arbitrary modeling decisions such as whether prices or wages respond first following an amenity improvement. Likewise, the equilibrium is conditional on the upper bound used to limit changes in the wage rate. Given these caveats, welfare measures

⁸² Recall from section III that “absentee shareholders” are defined as the owners of firms.

calculated using the results from the HEP-HEW algorithm should be interpreted as approximate and preliminary estimates of bounds on the true measures.

V. Simulating the Response to New 8-hour Ozone Standards

This section uses the general equilibrium framework to simulate how households and markets would respond if ozone concentrations in the San Francisco-Sacramento region were decreased to satisfy recently revised standards. Part (A) describes the regulatory background for the new limits on ozone concentrations. Much of the San Francisco-Sacramento region currently fails to meet these standards. In part (B) an “attainment” scenario is constructed to predict the community-specific ozone levels that would correspond to the hypothetical situation where the entire region satisfies the new standards. Then, four different versions of the general equilibrium model are used to analyze the welfare implications of the air quality improvement. Part (C) provides an overview of the simulations, part (D) reports results from vertically and horizontally differentiated versions of the single-market model, and (E) reports results from the dual-market model in the “free mobility” and “limited mobility” cases.

(A) Policy Background

Due to concern over the negative health effects of ground-level ozone, the Environmental Protection Agency (EPA) recently established stricter standards for ambient ozone concentrations. The *Phase 2 Ozone Rule* finalized in November, 2005 identifies areas that fail to meet the national 8-hour limit of 0.08 parts per million (ppm). These areas comprise 18% of the counties in the United States. The Phase 2 Rule outlines a set of corrective actions that must be taken by the states which contain these “non-attainment areas”. First, they are required to install abatement equipment that EPA considers to be reasonably available. Second, they must restrict construction of new stationary sources of ozone precursors and restrict modifications to existing sources⁸³. Finally, the states that contain

⁸³ Ozone forms when oxides of nitrogen (NOx) and reactive organic gasses (ROG) are combined with sunlight.

areas with the most severe violations may be required to use cleaner-burning reformulated gasoline.

California contains the four non-attainment areas with the most severe violations of the 8-hour limit: the Los Angeles Air Basin, the Sacramento CMSA, Riverside County, and the San Joaquin Valley. These are the only regions in the United States that EPA designates as having a “serious” or a “severe” problem. State officials addressed this problem in May, 2006 by adopting an even stricter 8-hour standard of 0.07 ppm. Meeting this standard will require a substantial reduction in emissions of ozone-forming chemicals.

To address air quality concerns in California and in other parts of the country, state and federal lawmakers have passed a series of regulations which have taken effect recently or will take effect in the near future. These include: stricter standards for emissions from cars, light trucks, and SUV’s (2004); limits on emissions from paint removers (2005); low sulfur levels in diesel fuel (2006); tighter emissions standards for heavy-duty diesel trucks (2007); greenhouse gas emissions standards for passenger cars and light trucks (2009); and tighter emissions standards for off-road diesel equipment (2011). The California Air Resources Board predicts that these regulations will lead to large decreases in emissions of ozone-forming chemicals by 2020, specifically emissions of oxides of nitrogen (NO_x) and reactive organic gasses (ROG). Within the San Francisco-Sacramento area, emissions of NO_x are predicted to decrease by 51% and emissions of ROG are predicted to decrease by 39% (Alexis and Cox [2005]).

The new ozone standards and emissions limits raise a number of questions for policy analysis. One possibility would be to analyze the welfare implications for the bundle of state and federal regulations on emissions. This would require translating the predicted decreases in NO_x and ROG into a corresponding reduction in ground level ozone. The difficulty with this idea is that predicting the concentration of ozone in a particular community would require modeling the formation of ozone, as well as wind patterns and other atmospheric conditions that influence how it moves within and across air basins. This is left as a possibility for future research. Instead, the rest of this section investigates the benefits of the new ozone standard itself. In other words, the analysis asks the hypothetical question: what

would households be willing to pay to reduce ozone just enough so that no community exceeds the 0.07 ppm limit?

(B) Developing the Attainment Scenario

Two issues must be addressed in order to assign an ozone measure to each community in the attainment scenario. First, the ozone-based proxy for air quality that was used during the estimation is measured differently than the 8-hr reading used to define California's new standard. Therefore, for the general equilibrium simulation to be theoretically consistent, "attainment" must be quantified in terms of the proxy measure used during the estimation. The second issue is that the implicit reductions in emissions needed to bring communities with high ozone levels into attainment would be likely to spill over into communities that already satisfy the new standards. For example, 8-hr ozone levels in the western part of the San Francisco Bay Area already fall below the new 0.07 ppm limit. This is partly due to natural wind currents that tend to push ozone toward the eastern side of the Bay and eventually into Sacramento. This ozone migration is part of the reason why Sacramento currently fails to meet the 8-hr standard. Assuming ozone reductions in the attainment scenario are derived from the type of regulations on emissions listed above, a reduction in ozone that would bring Sacramento into attainment would stem (at least partially) from reduced emissions of ozone precursors in the San Francisco Bay Area. Since wind currents do not remove ozone from the Bay Area entirely, the decrease in emissions would lower ozone levels there as well.

The attainment scenario was developed in two steps. First, data on 8-hr ozone levels were used to identify individual air quality monitoring stations that violated the 0.07 ppm limit during the years of the study period (1999-2001). For each of these "nonattainment" monitors, the reduction in 8-hr ozone required to bring it into attainment was translated into the corresponding reduction in the alternative (1-hour) ozone measure that was used to define air quality during the estimation. The second step was to estimate the historical relationship between changes in ozone at nonattainment stations and changes in ozone at attainment stations (i.e. those with 8-hr concentrations below 0.07 ppm during 1999-2001). The

resulting estimate was used together with the ozone reductions predicted for nonattainment stations to predict the corresponding reductions for attainment stations.

Chapter 4 described how the level of air quality in each community was measured using a three-year average of the annual top 30 1-hr daily maximum ozone readings. In contrast, California's new ozone standard is based on a 3-year average of the 4th highest annual 8-hr ozone reading. To make the conversion between the two measures, the reduced-form relationship between them was estimated using annual data for each monitoring station in the study region. First, the 8-hr measure was constructed for each monitoring station for 1999-2001. Then, the 1-hr data were regressed on the 8-hr data. The results imply that a 1 ppm reduction in the 8-hr measure corresponds to a 1.09 ppm reduction in the 1-hr measure. The coefficient was highly significant and explained almost all of the variation in ozone concentrations ($R^2 = .97$). Next, for all nonattainment monitors, the reduction in the 8-hr measure needed to bring that monitor into attainment was multiplied by 1.09 to generate an estimate for the corresponding reduction in the 1-hr measure.

During the second step of the process, annual data from 1980 to 2003 were used to predict the percentage change in ozone at the average nonattainment monitor that would correspond to a 1 ppb decrease in ozone at the average attainment monitor. More precisely, the annual change in the log of 8-hr ozone (ppb) averaged over all attainment monitors was regressed on the annual change in 8-hr ozone averaged over all nonattainment monitors. The results indicate that a 1 ppb decrease at the average nonattainment monitor has historically been accompanied by a 0.06% decrease at the average attainment monitor. However, this relationship was not statistically significant. While ozone concentrations have decreased substantially in the areas with the lowest quality air, they have remained relatively flat in attainment regions. Figure 6.3 displays this trend, using data from representative monitoring stations in Sacramento and Oakland. The figure illustrates how weather patterns near the San Francisco Bay limit the sensitivity of ozone concentrations to changes in precursor emissions. Yet it seems unlikely that large reductions in NO_x and ROG in Oakland would have no effect on ozone. Therefore, for the monitoring stations that were in attainment of the new standards during 1999-2001, the percentage ozone reduction in the attainment scenario

was estimated by multiplying the average change at nonattainment stations (-11.6 ppb) by the regression coefficient (0.06%). This calculation implies a 0.7% decrease. In comparison, ozone levels are predicted to decrease by an average of 12% at nonattainment stations.

Finally, the 1-hr ozone levels predicted for the individual monitoring stations were used to produce community-specific measures of air quality, using the same aggregation procedure outlined in chapter 4. Figure 6.4 shows the difference between the attainment scenario and baseline ozone concentrations in 1999-2001. Air quality improves significantly for most communities physically located in the Sacramento and Yolo primary metropolitan statistical areas, as well as for some communities in Oakland and San Jose. In contrast, the ozone reductions are marginal for communities in the San Francisco, Santa Cruz, and Santa Rosa PMSAs. Table 6.4 reports the average reduction over all the communities in each PMSA, as well as the range of changes across the individual communities. The reductions range from 0.6% in the San Mateo Union High school district (San Francisco PMSA) to 27.9% in the Black Oak Mine Unified school district (Sacramento PMSA).

(C) Overview of the Simulations

The general equilibrium framework was used to simulate how households and markets would adjust to the attainment scenario under each of the four equilibrium concepts discussed in section II. In each case, the population of households was defined using the joint distributions of income and preferences described in chapter 5 (using the uniform assumption in the horizontally differentiated cases). To reduce the computational burden, the simulations were based on a 1-in-10 sample of the population. Furthermore, the three versions of the simulation that depict households as having horizontally differentiated preferences were performed using a subset of communities. In total, 58 of the 122 communities were used: 8 from each of the Oakland, San Francisco, San Jose, Santa Rosa, Vallejo, and Sacramento PMSAs, and all 5 communities in each of the Santa Cruz and Yolo PMSAs. The communities in this subset span the range of values for air quality in the study region, and also provide some of the highest and lowest levels of school quality and $\bar{\xi}$.

As in the exploratory simulations, the highest possible wage that could be paid to

workers in each occupation was set to the maximum observed in the initial equilibrium. This effectively capped wages in the San Francisco and San Jose PMSAs which, together, account for highest wage observed for 18 of the 22 occupations. In the “free mobility” version of the simulation, the 58 communities translate to 124 possible (community, PMSA) locations. In the “limited mobility” simulation, households were restricted to live and work in the same PMSA. Thus, the simulation only used data on households who were observed making this choice in the initial equilibrium.

Finally, unlike the other three scenarios, the vertically differentiated simulation was conducted using all 122 communities. This is because reducing the choice set from 122 to 58 communities was found to produce a fivefold increase in the average WTP^{GE} —the opposite result from reducing the choice set when households have horizontally differentiated preferences (table 6.2). This result stems from the one-dimensional ordering of households across communities. Intuitively, removing a community from the middle of the price ordering requires the price of housing to fall in the next most expensive community⁸⁴. This produces a “domino effect” which systematically decreases the price of housing in all the subsequent communities. These price decreases dominate the resulting general equilibrium welfare measures.

Analysis of the results from the four simulations focuses on two questions. First, what is the pattern of results for capitalization and welfare effects? Second, how do those results vary with the equilibrium concept that underlies the simulation?

(D) *Simulation Results: Single-Market*

Table 6.5A reports the average changes in ozone, prices, and welfare for the attainment and nonattainment communities in each PMSA calculated using the results from the horizontally differentiated version of the model. All welfare measures are based on households’ initial locations. For example, the \$93 MWTP and \$612 WTP^{GE} for nonattainment communities in

⁸⁴ More formally, recall from section III that when communities are strictly ranked according to public goods provision, the single crossing property requires every household in community j to prefer j to $j+1, j+2, j+3, \dots, J$. Therefore, a household on the “border” of $j+2$ and $j+3$ must have a higher HEP for $j+3$ than every household living in $j+1$. When $j+2$ is removed from the choice set, the new “border” household between $j+1$ and $j+3$ will necessarily define a lower price for $j+3$.

Sacramento are based on the households who started in those communities, regardless of where they moved during the simulation. Not surprisingly, the table displays the type of sorting behavior that underlies the empirical model and the estimator. On average, households who initially chose to live in nonattainment communities have a lower MWTP for improved air quality (\$70) than households who chose to live in attainment communities (\$134). The reverse is true for partial equilibrium welfare measures because nonattainment communities have larger ozone reductions. These air quality improvements make the (formerly) nonattainment communities relatively more desirable, leading households to bid up their rental prices. The ensuing capitalization partly offsets welfare improvements from the ozone reduction.

Since the simulation framework is a “closed” model and there are no income effects, the 58 housing markets cannot clear simultaneously if housing prices increase everywhere. As more households move into nonattainment communities and the demand for housing decreases in their initial locations, some prices must fall. The communities with relatively small air quality improvements are generally the ones with price decreases, adding to the welfare improvement for current residents. Table 6.5B illustrates this pattern by showing results for the eight communities in the Oakland PMSA. As expected, the relatively large ozone reductions in Acalanes (-10.9%) and Livermore Valley (14.6%) are accompanied by price increases (0.7% and 1.0%). Likewise, the relatively small ozone reductions in Alameda City, Antioch, Berkeley, Castro Valley, and Dublin lead to small decreases in the price of housing. For example, although ozone concentrations drop by 0.7% in Alameda City, the much larger reductions in communities like Acalanes and Livermore Valley make them more attractive, inducing some of the households in Alameda City to move there. Consequently, the price of housing in Alameda City has to drop by -0.3% to clear the market. Albany City is an exception to this pattern. The 0.2% increase in its rental price of housing is somewhat surprising given its relatively small ozone reduction. A closer look reveals that Albany City provides a very high level of public goods overall. It ranks among the top 10 communities by school quality, air quality, and ξ ; the only community with higher values of all three public goods is 25% more expensive. This explains why a small ozone reduction induces

outsiders to bid up the price of housing. Antioch presents the opposite case. Out of the 58 communities, it ranks 39th by air quality, 41st by school quality, and 42nd by ξ . A moderate 2.1% ozone reduction is not enough to increase the demand for housing, given the improvements in other communities.

Table 6.5B also reveals some of the diversity in welfare effects by reporting how the average WTP^{GE} differs between households who move and households who do not. In Antioch and Dublin there is very little difference, whereas households who move out of Albany City, Alameda City, Berkeley, and Castro Valley generally have larger welfare improvements than non-movers. The opposite is true for the (Sacramento) communities in the last two rows of the table: Natomas and Black Oak Mine. Intuitively, some households move to take advantage of changes in the set of location choices, while others are effectively forced out by higher prices. The later case demonstrates how a policy that pursues “environmental justice” can make renters worse off. People who initially chose to live in nonattainment communities like Black Oak Mine and Acalanes may have done so because they are indifferent to air quality, or because they have low income. This type of household would prefer higher ozone levels and inexpensive housing to the higher prices that follow the ozone reduction. Consequently, their welfare goes down when prices increase. In Acalanes some households are able to cut their losses by moving, while in Black Oak Mine households who are forced out by the higher rents tend to have larger welfare decreases than non-movers.

When households are restricted to have vertically differentiated preferences, the simulation framework produces a similar pattern of price changes. To compare the results between the two simulations, tables 6.6A and 6.6B report price changes and welfare effects for the same 58 communities used to construct tables 6.5A and 6.5B. As in the horizontal case, equilibrium prices tend to increase (decrease) in (non)attainment communities following the ozone reduction. Notice that the vertical model depicts a much larger share of the population moving between communities (table 6.6B). In particular, every household in Antioch, Castro Valley, and Livermore Valley moves to a different community during the simulation. This reflects changes to the position of these three communities in the ranking by

overall public goods provision. For example, prior to the air quality improvement, Antioch was ranked 23rd by the public goods index. The ozone reduction decreases its rank to 21. While the ranking of communities changes, households are still strictly ordered across communities according to equation (6.6). The households who previously lived in Antioch are sorted into Travis and Vacaville, the new 22nd and 23rd ranked communities. The one-dimensional ordering of households and communities also leads to much less variation in the welfare obtained by movers relative to non-movers because it implies that they have similar preferences and occupy similar communities.

Although the vertical and horizontal simulations produce a similar pattern of price changes across attainment and nonattainment communities, the corresponding welfare measures are quite different. In the horizontal case, the welfare improvement for the average household in nonattainment communities (\$478) is twice as large as in attainment communities (\$236). Meanwhile, households in nonattainment communities have lower values for WTP^{GE} in the vertical model (\$201 compared to \$231). To illustrate the source of this difference, figure 6.5 shows how the average WTP^{GE} calculated from each of the two models varies across communities with the (pre-shock) air quality and price of housing. In the vertically differentiated model, communities with the highest initial ozone concentrations also have the lowest average WTP^{GE} . This is due to a combination of three factors. First, the communities with the lowest air quality are among the least expensive. Therefore the vertical model assigned them the lowest values for the public goods index during the estimation process, and the households who live there were assigned the lowest values for $\alpha | y$. These households will have the lowest per/unit WTP^{PE} for the change. Second, communities with higher initial ozone concentrations have larger reductions in the attainment scenario and, therefore, larger price increases. For many households in nonattainment communities, the welfare loss from the price increase outweighs the welfare gain from the ozone reduction. Finally, the strict ordering of households across communities means they have few opportunities to adjust to price increases by moving. On average, a household that moves during the simulation locates in a new community that differs from its initial location by 0.6 places in the price ranking. The combination of these three factors leads to a nearly

monotonic relationship between WTP^{GE} and the price of housing.

While the price increases in nonattainment communities are generally larger in the horizontal simulation, so is the welfare gain from the ozone reduction. This is because, during the estimation process, the horizontal model recognized that there may be some households in nonattainment communities who have relatively strong preferences for air quality and chose to live there for other reasons. In addition, since $\bar{\xi}$ is positively correlated with the price of housing, the estimator often assigns higher values for $\gamma_{air}/\gamma_{\bar{\xi}}$ to households in less expensive communities. This helps to explain the downward trend of WTP^{GE} in the price of housing. Finally, since households in the horizontal simulation are not strictly ordered across communities, they have more flexibility in how they adjust to the change. The average mover chooses a new location that differs from its initial community by 3.3 places in the price ranking.

(E) Simulation Results: Dual-Market

Table 6.7A reports the average changes in prices, wages, and welfare for each PMSA in the “free mobility” version of the dual-market simulation. As in the exploratory simulations, the results appear to be driven largely by a cycle of price and wage increases, resulting in higher housing prices for every community. For nonattainment communities, the average increase is 7.9% compared to 1.8% in attainment communities. When the price of housing increases, wages have to increase to ensure that the supply of workers continues to meet the demand. As a result, the PMSAs with the largest ozone reductions also tend to have the largest wage increases. This is one reason why average wages increase very little in San Jose (\$217) and San Francisco (\$220) which are within commuting distance to relatively few nonattainment communities. Of course, there is also little scope for wages to increase in those two PMSAs since they were used to set the maximum wage for 18 of the 22 occupations. Although communities in San Francisco and San Jose had relatively small price increases, the increases are still sufficiently large that the average household experienced a welfare loss.

Of the 94,594 working households in the simulation, 14,012 moved to a different

community but only 464 changed their job location. This is not surprising given that $\theta_1 < 1$ for 99% of working households, which means they would be paid less than the average wage in any job location other than their initial one. This reinforces the interpretation of θ_1 as a “job mobility” parameter. Table 6.7B further illustrates the relationship between job mobility, wage adjustments, and welfare effects. In general, workers who are more mobile have larger increases in salary following the ozone reduction. For example, of the eight occupations in the table, workers in *management, computer and mathematical*, and *architecture and engineering* occupations have the largest increases in salary and the highest average values for the job mobility parameter. Moreover, the top 1% of workers in each of these occupations has values for θ_1 above 1. As these highly mobile workers move to different job locations, average salaries increase in their original PMSAs, increasing wage income and welfare for the workers who do not move. This pattern is reversed for workers in *healthcare support, personal care and service*, and *farming, fishing, and forestry* occupations. Low values for θ_1 effectively limit their ability to react to the higher housing prices by moving to a new job location. Consequently, adjustments to the wage rate are relatively small. Overall, salaries increase for 90 of the 175 (occupation, PMSA) combinations, remain unchanged for 43 and decrease for the remaining 42.

Restricting households to live and work in the same PMSA reduces the magnitude of the price and wage increases. This can be seen in table 6.8, which reports average changes in prices, wages, and welfare for this “limited mobility” version of the simulation. Notice that the summary statistics for the partial equilibrium welfare measures differ slightly from the ones reported in table 6.7A. This is because the simulation excluded workers who were observed commuting between PMSAs in the initial equilibrium, reducing the sample size by approximately 15% and reducing the number of possible (community, PMSA) combinations from 124 to 66. The impact on welfare measures is most pronounced in the attainment communities where smaller price increases cause the average WTP^{GE} to roughly triple relative to the unrestricted simulation.

Removing the possibility for between-PMSA commuting partially decouples housing prices in a given PMSA from wage adjustments in other labor markets. This has the largest

effect on communities in the Yolo PMSA, where the average increase in housing prices following the ozone reduction drops from 8.86% in the free mobility simulation to 4.19% in the limited mobility case. This is because 24.6% of Yolo's working population commutes to work in the Sacramento PMSA (table 4.2). In the free mobility simulation, the increase in wages paid to these workers increases their demand for housing in Yolo, which requires the price of housing to increase. Similar connections between the housing communities in Vallejo and labor markets in Sacramento, Yolo, and Oakland help to explain why the price increases in Vallejo's communities are much smaller in the limited mobility simulation. Of course, households who live in Sacramento and are paid higher wages may still bid up the price of housing in Vallejo and Yolo, which is why limits on commuting do not fully decouple the relationship between prices in one PMSA and wages in another.

VI. Conclusions

This chapter developed a framework to simulate general equilibrium adjustment to changes in local public goods and then used the framework to investigate the welfare implications for improving air quality in California's San Francisco-Sacramento region. Alternate versions of the framework were developed to depict different equilibrium concepts, each of which is nested by the idea of a dual-market locational equilibrium. Modeling households as having horizontally differentiated preferences increased the computational burden of the simulations, and allowing labor market adjustment raised the possibility of divergent solutions and multiple equilibria. Strategies for addressing both issues were developed and tested. Preliminary evidence suggested that using a subset of communities and a fraction of the population can reduce the computational burden but will tend to understate welfare measures in the single-market case. In the dual-market case, imposing a limit on the maximum wage allowed the model to converge but welfare effects appeared to be dominated by a cycle of price and wage increases. This cycle essentially transfers wealth from retired households and absentee shareholders to working households and absentee homeowners. To the extent that households are homeowners rather than renters, this will overstate (understate) the true

willingness-to-pay for working (retired) households. Given these caveats, the framework was used to simulate how households and markets would respond if emissions of ozone precursors were reduced just enough for the entire San Francisco-Sacramento region to be in attainment of California's new 8-hour standard for ozone.

Table 6.9 summarizes results for each version of the model by reporting the average per/household change in rents, wages, and welfare. In the vertical model, accounting for general equilibrium adjustment produced an estimate for the average willingness-to-pay of \$229 compared to a partial equilibrium measure of \$130. The difference between them is mainly due to a \$98 decrease in annualized expenditures on housing (i.e. rents). When households were depicted as having horizontally differentiated preferences, there was little difference between partial equilibrium (\$348) and general equilibrium (\$335) welfare measures, although both exceeded the vertical estimates. Allowing labor markets to adjust led to general equilibrium welfare measures (e.g. \$879) that were more than double their partial equilibrium counterparts (e.g. \$393). However, this result appeared to be dominated by the cycle of price and wage increases. Limiting commuting possibilities decreased the magnitude of this cycle.

Approximately 60% of the houses in the San Francisco-Sacramento region are owner occupied. If the simulation framework were calibrated to reproduce this share, the cycle of price and wage increases would probably have less of an impact on general equilibrium welfare measures. This hypothesis is left as a possibility for future research. A second way to extend the model would be to explicitly introduce moving costs. While the job mobility parameter already acts as a de facto moving cost for working households, it is still costless to move between communities within a PMSA. Chapter 7 considers the implications of this extension.

Finally, assuming the "cycle effect" on dual-market welfare measures outweighs the "sampling" effect, the \$879 estimate could be interpreted as an upper bound on the ex post willingness-to-pay for the air quality improvement. Likewise, the \$335 estimate from the single-market version of the model could be treated as a lower bound. Multiplying these bounds by the number of households in the study region would imply cumulative annualized

benefits between \$1.1 and \$2.8 billion.

Table 6.1: Sensitivity of the Simulation Results to Using a Fraction of the Sample Size

Community	Ozone (ppb)			Average MWTP (1 ppb reduction)			Average WTP ^{PE}			Households who Move (%)			Change in Price Index (%)			Average WTP ^{GE}		
	Initial Level	New Level	Change	1%	10%	100%	1%	10%	100%	1%	10%	100%	1%	10%	100%	1%	10%	100%
Acalanes	86	76	-12%	17	17	17	182	181	181	2.8	2.5	2.4	1.3	0.5	0.4	-355	-16	28
Alameda City	31	31	0	399	395	395	0	0	0	0.7	0.4	0.4	0.0	-0.5	-0.5	-18	182	191
Albany City	54	54	0	222	217	221	0	0	0	1.4	0.5	0.6	0.5	-0.2	-0.2	-209	68	77
Alexander Valley	71	71	0	66	73	73	0	0	0	5.9	0.8	0.5	0.0	-1.0	-1.1	1	346	366
Alum Rock	66	66	0	73	72	73	0	0	0	0.0	0.4	0.4	0.1	-0.9	-0.9	-33	288	305
Antioch	78	76	-3%	68	68	68	139	139	139	0.0	0.1	0.2	0.9	-0.1	-0.4	-67	162	240
Benicia	37	37	0	357	368	370	0	0	0	4.5	3.2	3.3	0.3	-1.4	-1.6	-62	478	546
Berkeley	47	47	0	217	220	220	0	0	0	0.8	0.3	0.3	0.2	-0.3	-0.3	-62	104	110
Berryessa	66	66	0	81	82	82	0	0	0	2.4	2.6	2.4	0.2	-0.8	-0.9	-85	346	388
Black Oak Mine	106	76	-28%	79	63	68	2,492	2,064	2,205	4.4	4.5	4.8	27.3	22.2	21.9	-2,130	-1,749	-1,532
Cabrillo	52	52	0	159	161	162	0	0	0	1.3	0.5	0.5	0.2	-0.3	-0.3	-89	145	154
Calistoga	60	60	0	105	103	104	0	0	0	0.0	4.6	3.5	0.1	-1.0	-1.1	-24	283	307
Cambrian	72	72	0	44	44	45	0	0	0	10.5	5.5	6.5	0.4	-0.2	-0.2	-187	77	104
Campbell	71	71	0	30	32	32	0	0	0	0.0	0.1	0.1	0.2	-0.1	-0.2	-66	58	67
Castro Valley	63	63	0	122	122	122	0	0	0	3.4	1.8	1.9	0.0	-1.0	-1.1	4	377	398
Center	94	76	-19%	140	137	138	2,513	2,468	2,489	0.0	0.0	0.03	14.5	9.8	9.6	-140	652	723
Cloverdale	71	71	0	66	63	63	0	0	0	0.0	2.1	4.0	0.5	-1.8	-2.1	-103	400	460
Cotati-Rohnert Park	60	60	0	98	98	98	0	0	0	0.0	0.8	0.7	0.0	-1.4	-1.6	0	351	403
Cupertino	52	52	0	84	83	84	0	0	0	0.0	0.04	0.05	0.2	-0.1	-0.1	-124	34	39
Davis	87	76	-13%	21	23	23	243	268	267	1.9	1.9	1.8	2.4	1.2	1.1	-415	-50	-21

Table 6.2: Sensitivity of the Simulation Results to Using a Subset of Communities
(1-in-100 sample of the population)

Community	% ozone reduction	Average MWTP (1 ppb)	Average WTP ^{PE}	Price change (%)		Movers (%)		WTP ^{GE}	
				J = 20	J = 122	J = 20	J = 122	J = 20	J = 122
Acalanes	-12%	17	179	1.3	1.0	2.8	24.1	-355	-246
Alameda City	0	399	0	0.0	0.0	0.7	6.2	-18	-10
Albany City	0	222	0	0.5	0.5	1.4	1.4	-209	-203
Alexander Valley	0	66	0	0.0	-0.1	5.9	94.1	1	74
Alum Rock	0	73	0	0.1	0.0	0.0	5.7	-33	13
Antioch	-3%	68	94	0.9	-1.6	0.0	43.9	-67	490
Benicia	0	357	0	0.3	-0.4	4.5	16.4	-62	304
Berkeley	0	217	0	0.2	0.1	0.8	1.9	-62	-34
Berryessa	0	81	0	0.2	0.0	2.4	6.6	-85	5
Black Oak Mine	-28%	79	1,876	27.3	18.3	4.4	40.0	-2,130	-1,211
Cabrillo	0	159	0	0.2	0.1	1.3	3.8	-89	-37
Calistoga	0	105	0	0.1	-0.1	0.0	16.2	-24	40
Cambrian	0	44	0	0.4	0.2	10.5	9.6	-187	-74
Campbell	0	30	0	0.2	0.1	0.0	0.6	-66	-43
Castro Valley	0	122	0	0.0	-0.1	3.4	7.8	4	32
Center	-19%	140	2,887	14.5	8.0	0.0	100.0	-140	1,111
Cloverdale	0	66	0	0.5	-1.1	0.0	48.3	-103	271
Cotati-Rohnert Park	0	98	0	0.0	-0.5	0.0	10.0	0	141
Cupertino	0	84	0	0.2	0.1	0.0	0.4	-124	-81
Davis	-13%	21	1,752	2.4	2.1	1.9	3.3	-415	-343

Table 6.3A: Sensitivity of Equilibria to the Order of Adjustment in the HEP-HEW Algorithm

Community	% ozone reduction	Average MWTP (1 ppb)	Average WTP ^{PE}	Price change (%)		Movers (%)		Average WTP ^{GE}	
				rents first	wages first	rents first	wages first	rents first	wages first
Acalanes	-12%	17	179	0.57	0.62	2.67	2.82	365	369
Alameda City	0	387	0	0.41	0.47	1.29	1.40	276	281
Albany City	0	252	0	1.58	1.67	9.76	10.58	-226	-238
Alexander Valley	0	54	0	0.28	0.36	3.82	3.05	583	583
Alum Rock	0	170	0	-0.27	-0.17	1.32	1.32	152	122
Antioch	-3%	47	94	0.07	0.18	1.52	1.68	522	527
Benicia	0	318	0	2.48	2.72	2.01	2.24	-162	-199
Berkeley	0	227	0	0.20	0.24	0.57	0.59	212	213
Berryessa	0	121	0	-0.38	-0.32	2.72	2.76	179	154
Black Oak Mine	-28%	56	1,876	32.24	32.81	35.59	35.71	584	582
Cabrillo	0	150	0	0.14	0.18	1.25	1.18	-66	-87
Calistoga	0	133	0	-0.07	0.02	11.23	9.77	546	544
Cambrian	0	52	0	0.14	0.19	5.32	5.52	-63	-82
Campbell	0	30	0	0.14	0.18	0.65	0.67	-31	-45
Castro Valley	0	69	0	0.60	0.66	2.38	2.44	200	208
Center	-19%	169	2,887	19.03	19.54	17.35	17.50	3,524	3,521
Cloverdale	0	58	0	1.29	1.37	6.59	6.89	122	115
Cotati-Rohnert Park	0	213	0	1.01	1.07	0.53	0.56	297	307
Cupertino	0	246	0	0.21	0.24	0.22	0.23	-96	-108
Davis	-13%	167	1,752	3.36	3.43	0.18	0.18	1,795	1,773

Table 6.3B: Min, Max, and Mean Changes in Wages, by PMSA (for all 22 occupational categories)

PMSA	Min		PMSA	Max		PMSA	Mean	
	rents first	wages first		rents first	wages first		rents first	wages first
Oakland	0	0	Oakland	2,566	2,785	Oakland	415	441
Sacramento	0	-1,024	Sacramento	9,004	9,211	Sacramento	2,723	3,322
San Francisco	0	-96	San Francisco	74	94	San Francisco	2	8
San Jose	0	0	San Jose	138	137	San Jose	13	11
Santa Rosa	0	0	Santa Rosa	2,778	2,763	Santa Rosa	236	504
Vallejo	0	0	Vallejo	2,419	2,535	Vallejo	283	194
Yolo	0	-255	Yolo	440	603	Yolo	31	28

Table 6.3C: Comparing the Welfare of Working and Retired Households*

Community	% ozone reduction	Average MWTP (1 ppb)	Average WTP ^{PE}	Price change (%)	Average WTP ^{GE}	
					Workers	Retired
Acalanes	-12%	17	179	0.62	618	-40
Alameda City	0	387	0	0.47	450	-69
Albany City	0	252	0	1.67	-8	-316
Alexander Valley	0	54	0	0.36	1,907	-101
Alum Rock	0	170	0	-0.17	175	18
Antioch	-3%	47	94	0.18	685	63
Benicia	0	318	0	2.72	-80	-471
Berkeley	0	227	0	0.24	429	-12
Berryessa	0	121	0	-0.32	192	95
Black Oak Mine	-28%	56	1,876	32.81	2,486	-2,600
Cabrillo	0	150	0	0.18	-51	-25
Calistoga	0	133	0	0.02	1,204	-14
Cambrian	0	52	0	0.19	-73	-29
Campbell	0	30	0	0.18	-31	-34
Castro Valley	0	69	0	0.66	281	-148
Center	-19%	169	2,887	19.54	4,211	-717
Cloverdale	0	58	0	1.37	872	-204
Cotati-Rohnert Park	0	213	0	1.07	545	-148
Cupertino	0	246	0	0.24	57	-37
Davis	-13%	167	1,752	3.43	2,501	-409

* Welfare measures were calculated from the version of the algorithm in which wages adjust first.

Table 6.4: Summary of the Attainment Scenario, by PMSA

Primary Metropolitan Statistical Area	Average* of Top 30 1-hr Ozone			Changes in Individual Communities			
	Baseline (ppb)	Attainment (ppb)	Change (%)	communities in PMSA (#)	nonattainment communities (#)	largest reduction (%)	smallest reduction (%)
Oakland	77.0	71.1	-7.7	25	10	-14.6	-0.7
Sacramento	93.9	75.3	-19.7	16	15	-27.9	-0.8
San Francisco	55.6	55.2	-0.7	21	0	-0.8	-0.6
San Jose	74.2	70.8	-4.6	28	2	-8.4	-0.7
Santa Cruz	63.3	62.9	-0.7	5	0	-0.8	-0.7
Santa Rosa	65.1	64.6	-0.7	12	0	-0.8	-0.7
Vallejo-Fairfield-Napa	74.2	70.1	-5.5	10	5	-12.0	-0.7
Yolo	86.4	76.4	-11.6	5	5	-12.4	-8.2

* Calculated as an average over the communities in the PMSA, weighted by acreage.

**Table 6.5A: Summary of Average Welfare Effects and Price Changes, by PMSA
(Single-Market Horizontal)**

Primary Metropolitan Statistical Area	Communities in simulation (#)	Mean Δ Ozone *	Average Willingness-to-Pay for Ozone Reduction**							
			Non-Attainment Communities				Attainment Communities			
			Δ price (%)	MWTP	WTP ^{PE}	WTP ^{GE}	Δ price (%)	MWTP	WTP ^{PE}	WTP ^{GE}
Oakland	8	-7.7	-0.18	35	193	154	-0.31	253	75	193
Sacramento	8	-19.7	3.53	93	1,251	612	--	--	--	--
San Francisco	8	-0.7	--	--	--	--	-0.17	164	58	124
San Jose	8	-4.6	0.45	29	206	68	-0.36	69	32	165
Santa Cruz	5	-0.7	--	--	--	--	-0.38	80	37	171
Santa Rosa	8	-0.7	--	--	--	--	-1.26	101	44	380
Vallejo-Fairfield-Napa	8	-5.5	-1.88	72	273	729	-1.75	151	61	615
Yolo	5	-11.6	0.22	23	252	173	--	--	--	--
All Communities	58	-9.7	1.62	70	776	478	-0.58	134	50	236

* Weighted by each community's acreage. ** Weighted by population

**Table 6.5B: Average Welfare Effects and Price Changes for Individual Communities
(Single-Market Horizontal)**

Community	PMSA	Ozone (ppb)			Δ price (%)	movers (%)	MWTP	WTP ^{PE}	WTP ^{GE}	
		Baseline	Attainment	% Δ					movers	nonmovers
Acalanes	Oakland	85.8	76.4	-10.9	0.7	3.5	17	175	-36	-102
Alameda City	Oakland	31.2	31.0	-0.7	-0.3	2.8	395	84	598	179
Albany City	Oakland	54	53.8	-0.8	0.2	1.8	217	92	167	16
Antioch	Oakland	78.0	76.4	-2.1	-2.7	47.8	68	111	721	752
Berkeley	Oakland	47.0	46.6	-0.7	-0.2	0.9	220	77	247	149
Castro Valley	Oakland	63.2	62.8	-0.7	-0.8	3.9	122	52	431	330
Dublin	Oakland	75.0	71.5	-4.7	-0.3	8.6	68	242	368	362
Livermore Valley	Oakland	89.5	76.4	-14.6	1.0	7.8	19	287	-41	-94
Black Oak Mine	Sacramento	105.9	76.4	-27.9	20.9	37.8	63	2,049	-1,865	-1,216
Natomas	Sacramento	86.2	76.4	-11.4	1.2	15.2	77	774	466	600

**Table 6.6A: Summary of Average Welfare Effects and Price Changes, by PMSA
(Single-Market Vertical)**

Primary Metropolitan Statistical Area	Communities in simulation (#)	Mean Δ Ozone *	Average Willingness-to-Pay for Ozone Reduction**							
			Non-Attainment Communities				Attainment Communities			
			Δ price (%)	MWTP	WTP ^{PE}	WTP ^{GE}	Δ price (%)	MWTP	WTP ^{PE}	WTP ^{GE}
Oakland	8	-7.7	-0.02	38	212	257	-0.54	145	67	254
Sacramento	8	-19.7	2.33	24	309	201	--	--	--	--
San Francisco	8	-0.7	--	--	--	--	-0.47	122	42	206
San Jose	8	-4.6	-0.09	58	26	-21	-0.55	72	33	197
Santa Cruz	5	-0.7	--	--	--	--	-0.56	69	31	272
Santa Rosa	8	-0.7	--	--	--	--	-0.75	50	46	305
Vallejo-Fairfield-Napa	8	-5.5	-0.19	37	227	126	-0.80	66	28	211
Yolo	5	-11.6	0.40	49	59	243	--	--	--	--
All Communities	58	-9.7	1.25	32	245	201	-0.58	92	41	231

* Weighted by each community's acreage. ** Weighted by population

**Table 6.6B: Average Welfare Effects and Price Changes in Individual Communities
(Single-Market Vertical)**

Community	PMSA	Ozone (ppb)			Δ price (%)	movers (%)	MWTP	WTP ^{PE}	WTP ^{GE}	
		Baseline	Attainment	% Δ					movers	nonmovers
Acalanes	Oakland	85.8	76.4	-10.9	0.2	5.1	32	328	249	260
Alameda City	Oakland	31.2	31.0	-0.7	-0.5	7.7	246	51	263	264
Albany City	Oakland	54	53.8	-0.8	-0.5	17.5	101	42	283	281
Antioch	Oakland	78.0	76.4	-2.1	-0.8	100.0	36	263	229	--
Berkeley	Oakland	47.0	46.6	-0.7	-0.5	3.2	134	47	259	285
Castro Valley	Oakland	63.2	62.8	-0.7	-0.7	100.0	39	142	248	--
Dublin	Oakland	75.0	71.5	-4.7	-0.3	28.1	54	23	239	244
Livermore Valley	Oakland	89.5	76.4	-14.6	0.6	100.0	46	22	244	--
Black Oak Mine	Sacramento	105.9	76.4	-27.9	5.1	16.3	14	552	-128	-114
Natomas	Sacramento	86.2	76.4	-11.4	1.7	29.1	17	505	-33	-15

**Table 6.7A: Average Welfare Effects and Price Changes, by PMSA
(Dual-Market Free Mobility)**

Primary Metropolitan Statistical Area	Mean Δ Ozone * (%)	Mean Δ wages	Average Willingness-to-Pay for Ozone Reduction**							
			Non-Attainment Communities				Attainment Communities			
			Δ price (%)	MWTP	WTP ^{PE}	WTP ^{GE}	Δ price (%)	MWTP	WTP ^{PE}	WTP ^{GE}
Oakland	-7.7	1,890	3.31	27	162	173	2.13	246	72	252
Sacramento	-19.7	5,673	10.29	94	1,248	3,277	--	--	--	--
San Francisco	-0.7	220	--	--	--	--	1.00	158	55	-189
San Jose	-4.6	217	5.30	55	382	-541	1.51	131	60	-318
Santa Cruz	-0.7	1,065	--	--	--	--	1.43	212	97	219
Santa Rosa	-0.7	1,309	--	--	--	--	2.50	183	79	211
Vallejo-Fairfield-Napa	-5.5	3,156	5.32	59	197	745	4.05	134	54	976
Yolo	-11.6	2,422	8.86	125	1,211	1,602	--	--	--	--
All 58 Communities	-9.7	2,111	7.94	77	861	2,063	1.82	168	66	51

**Table 6.7B: Average Changes in Salary and Welfare by Occupation and Job Location
(Dual-Market, Free Mobility)**

Occupational Category (Standard Occupational Classification System)	Primary Metropolitan Statistical Area								Job Mobility (θ_1)	
	Oakland	Sacramento	San Francisco	San Jose	Santa Cruz	Santa Rosa	Vallejo-Fairfield-Napa	Yolo	mean	99th percentile
	<u>Change in Average Salary</u>									
Management	3,675	9,133	556	0	4,361	2,880	7,693	9,729	0.45	1.05
Computer & Mathematical	1,014	4,643	363	0	310	2,604	3,828	-39	0.47	1.04
Architecture and Engineering	897	4,943	0	0	777	3,011	4,467	152	0.48	1.07
Arts, Design, Entertainment, Sports and Media	1,642	5,771	0	0	0	217	1,240	-217	0.44	0.96
Healthcare Practitioners and Technicians	778	4,952	263	0	0	1,570	2,736	740	0.45	0.96
Healthcare Support	0	352	0	0	-68	0	188	-349	0.44	0.92
Personal Care and Service	0	42	0	0	-22	0	0	-139	0.39	0.94
Farming, Fishing and Forestry	0	108	-2	-9	-62	0	0	0	0.28	0.85
	<u>Average WTP^{GE}</u>									
Management	2,111	7,686	-289	-858	3,569	1,945	5,102	8,202	0.45	1.05
Computer & Mathematical	-63	3,691	-295	-668	-285	1,734	2,567	770	0.47	1.04
Architecture and Engineering	-291	3,805	-572	-700	168	2,121	3,090	-546	0.48	1.07
Arts, Design, Entertainment, Sports and Media	793	4,896	-405	-572	-416	-381	256	194	0.44	0.96
Healthcare Practitioners and Technicians	-343	4,099	-232	-680	-564	855	1,461	472	0.45	0.96
Healthcare Support	-354	68	-198	-245	-150	-270	-281	370	0.44	0.92
Personal Care and Service	-350	-360	-197	-306	-104	-347	-546	634	0.39	0.94
Farming, Fishing and Forestry		-360	-246	-445	-181	-378	-788	51	0.28	0.85

**Table 6.8: Average Welfare Effects and Price Changes, by PMSA
(Dual-Market, Limited Mobility)**

Primary Metropolitan Statistical Area	Mean Δ Ozone * (%)	Mean Δ wages	Average Willingness-to-Pay for Ozone Reduction**							
			Non-Attainment Communities				Attainment Communities			
			Δ price (%)	MWTP	WTP ^{PE}	WTP ^{GE}	Δ price (%)	MWTP	WTP ^{PE}	WTP ^{GE}
Oakland	-7.7	1,377	2.36	25	154	286	1.89	229	66	262
Sacramento	-19.7	4,493	9.47	94	1,257	2,691	--	--	--	--
San Francisco	-0.7	94	--	--	--	--	0.72	152	53	-177
San Jose	-4.6	113	4.62	56	385	-482	0.98	133	61	-188
Santa Cruz	-0.7	1,257	--	--	--	--	1.11	233	106	581
Santa Rosa	-0.7	1,231	--	--	--	--	1.28	190	82	544
Vallejo-Fairfield-Napa	-5.5	2,778	2.53	72	216	1,212	2.87	121	50	1,112
Yolo	-11.6	808	4.19	124	1,171	744	--	--	--	--
All 58 Communities	-9.7	1,809	6.79	80	902	1,827	1.24	165	65	165

Table 6.9: Average Per/Household Change in Welfare, Rents, and Wages by Equilibrium Concept

Locational Equilibrium Concept	WTP ^{PE}	WTP ^{GE}	Δ Rents*	Δ Wages
Single-market, vertical	130	229	-98	
Single-market, horizontal	348	335	30	
Dual market	393	879	1,138	2,111
<i>Retired households</i>	219	-422		
<i>Working households</i>	449	1,321		
Dual market, limited mobility	415	860	844	1,809
<i>Retired households</i>	219	-268		
<i>Working households</i>	484	1,310		

* Measured as housing expenditures per household

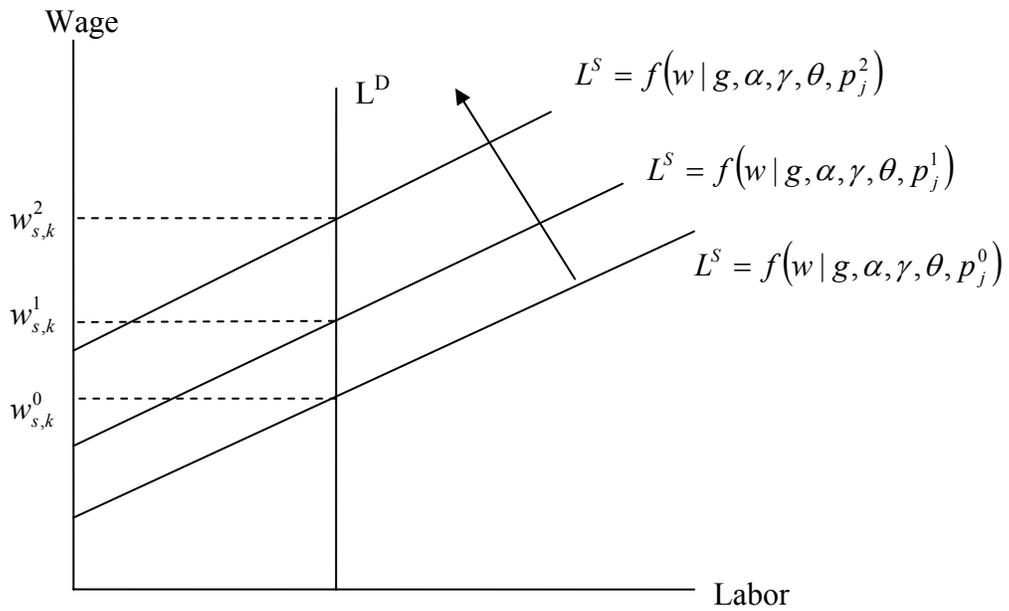
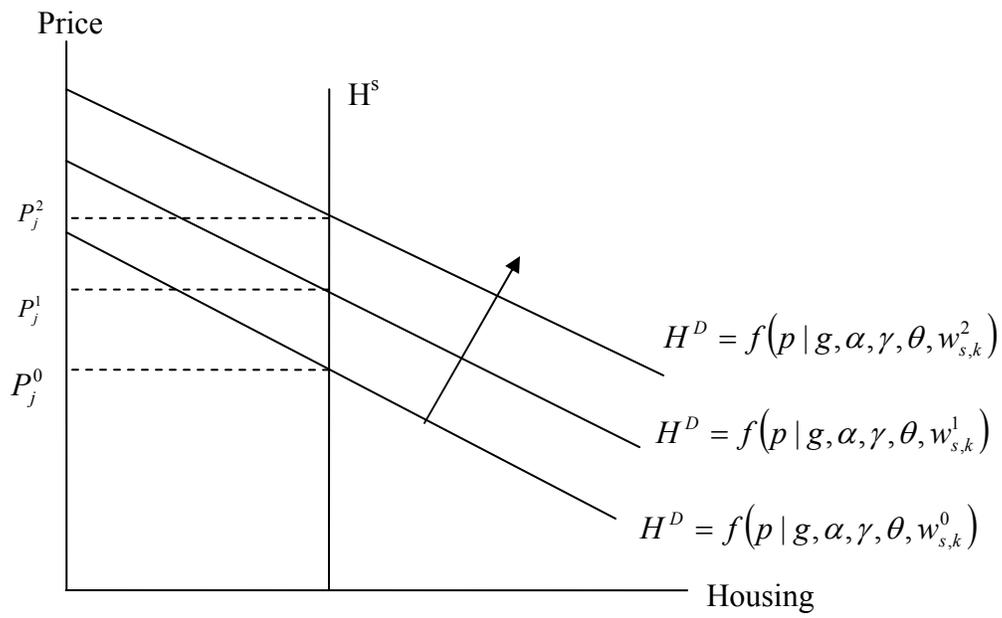


Figure 6.1: A Divergent General Equilibrium Solution

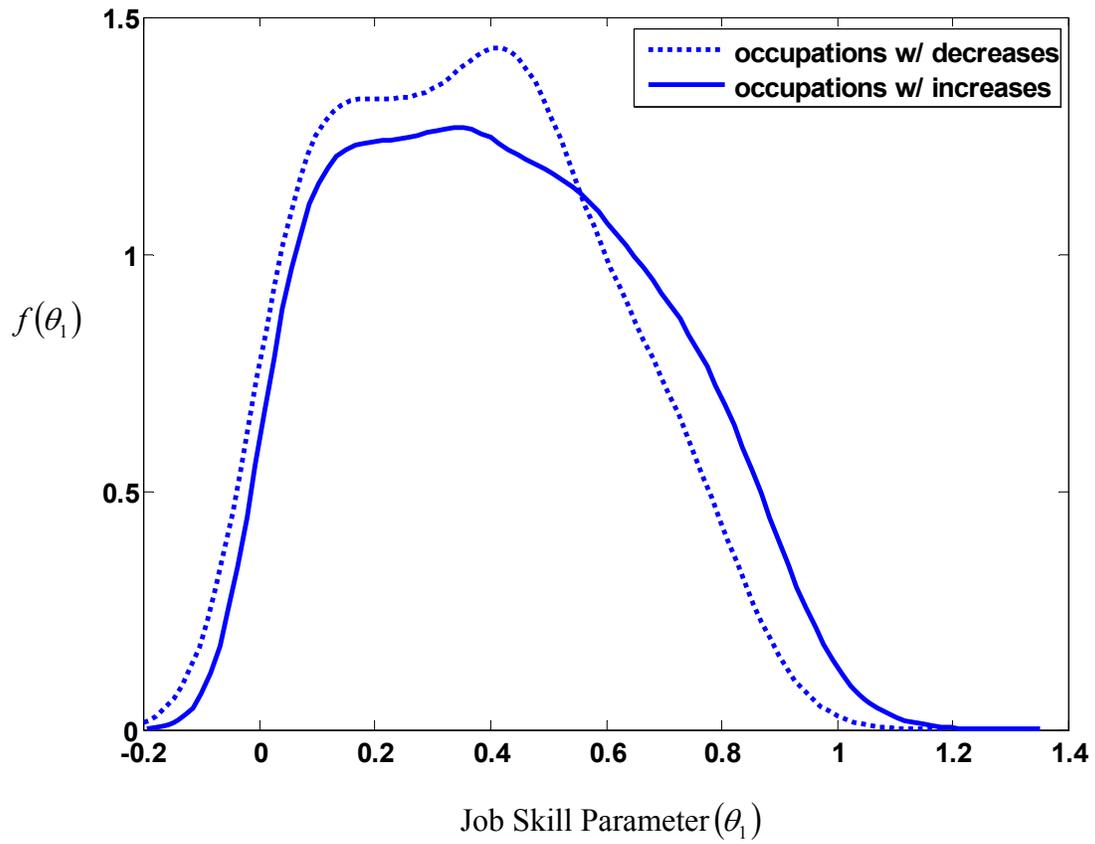


Figure 6.2: Distribution of Labor Market Mobility

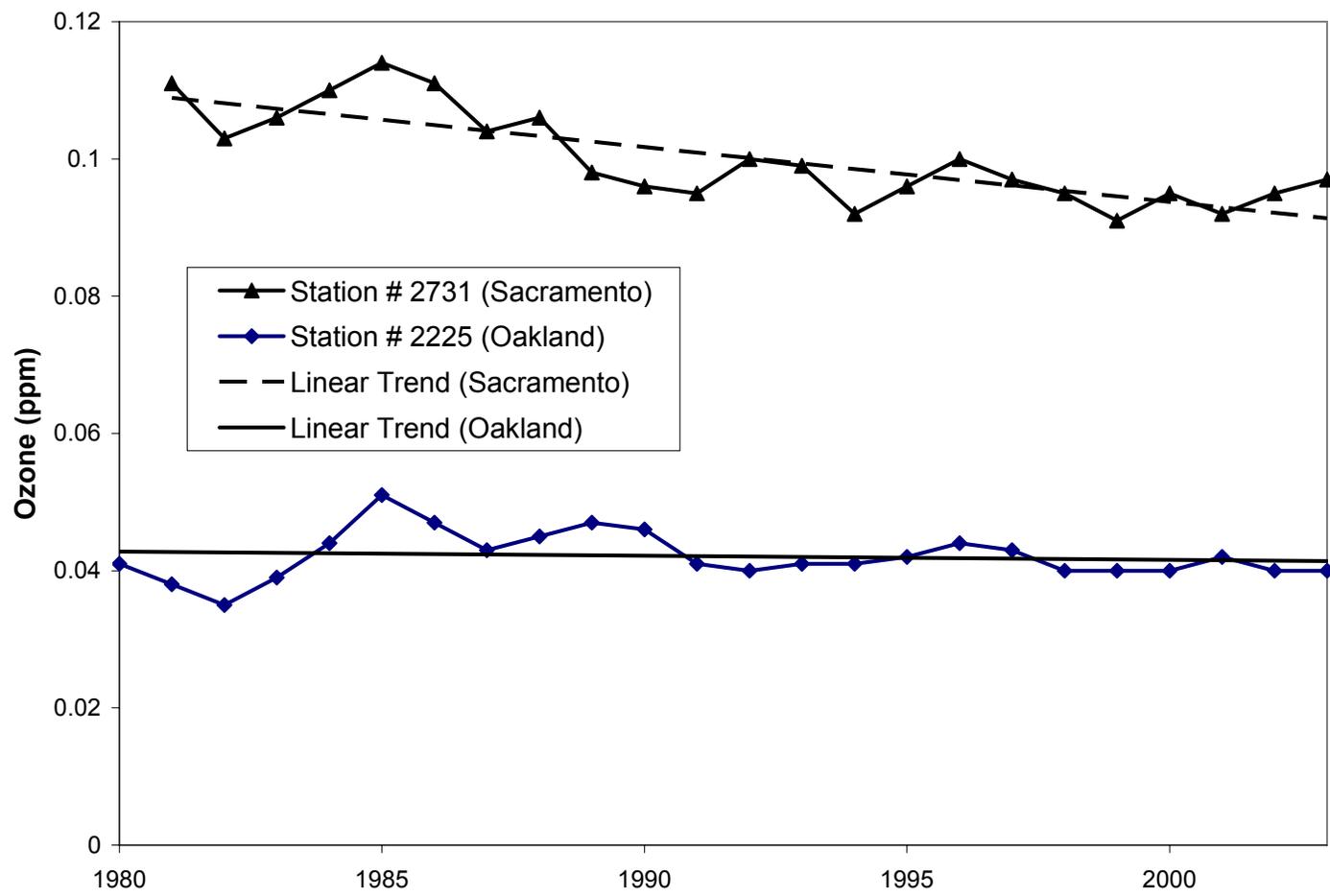


Figure 6.3: Ozone Concentrations in San Francisco and Sacramento, 1980-2003

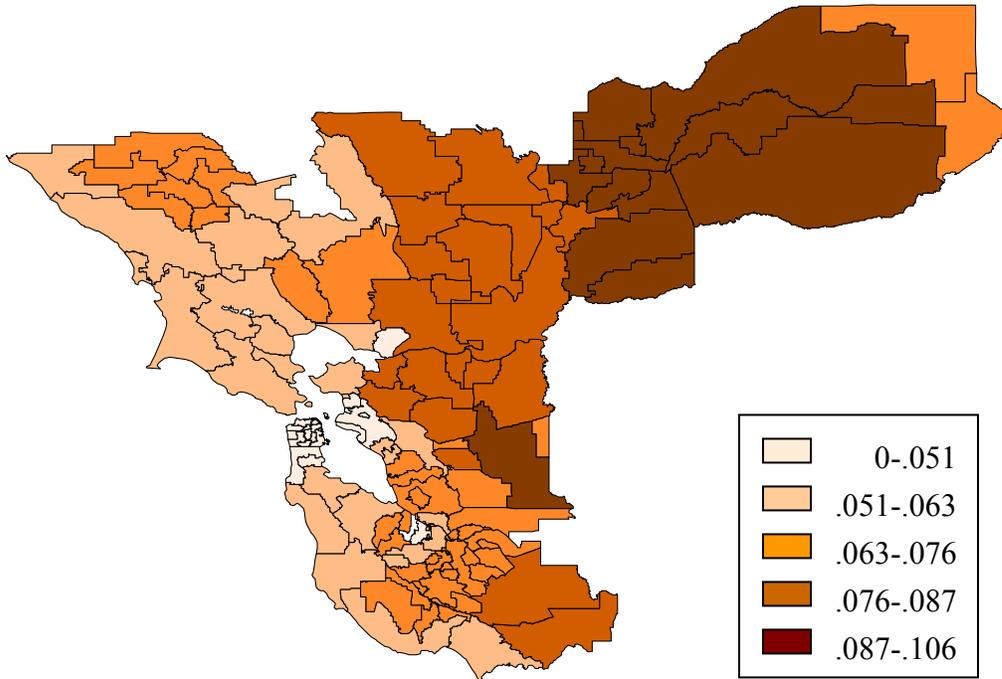


Figure 6.4A: Ozone (ppm) in the San Francisco-Sacramento Region, 1999-2001

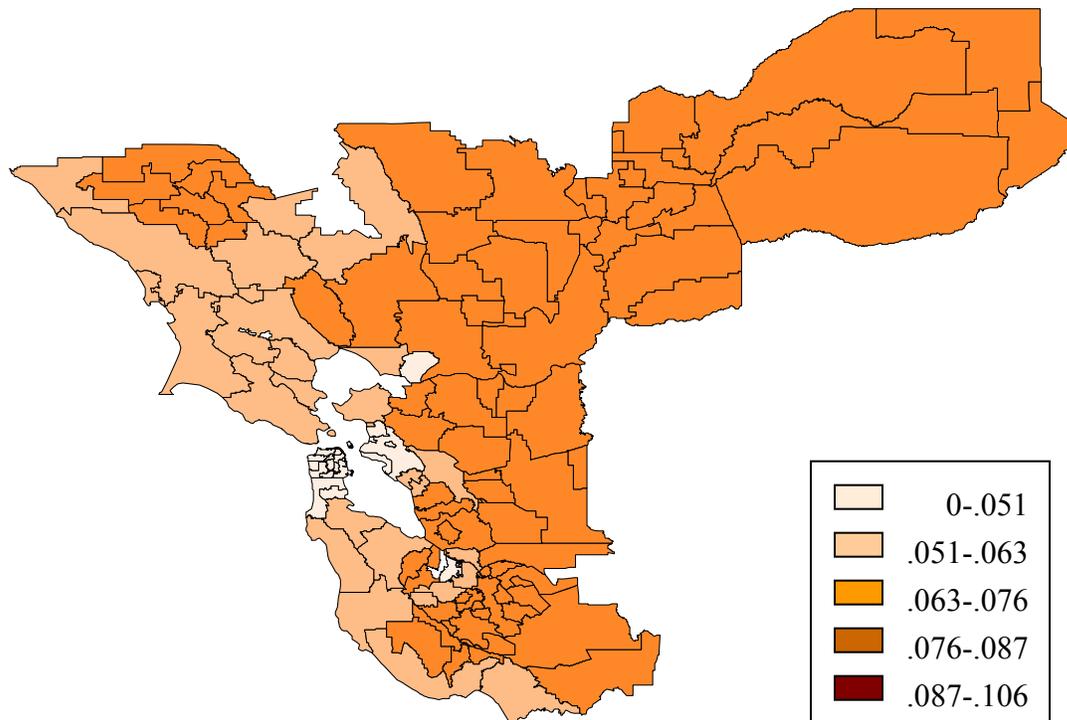


Figure 6.4B: The Study Region in Attainment of the 8-Hour Ozone Standard (0.07 ppm)

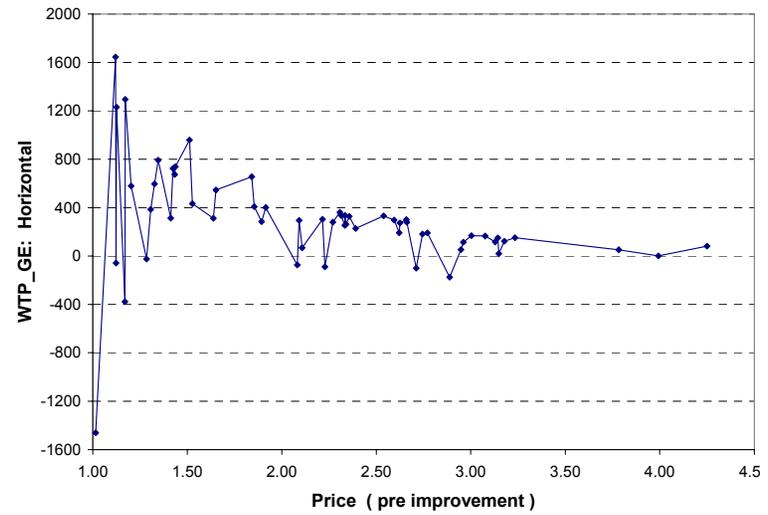
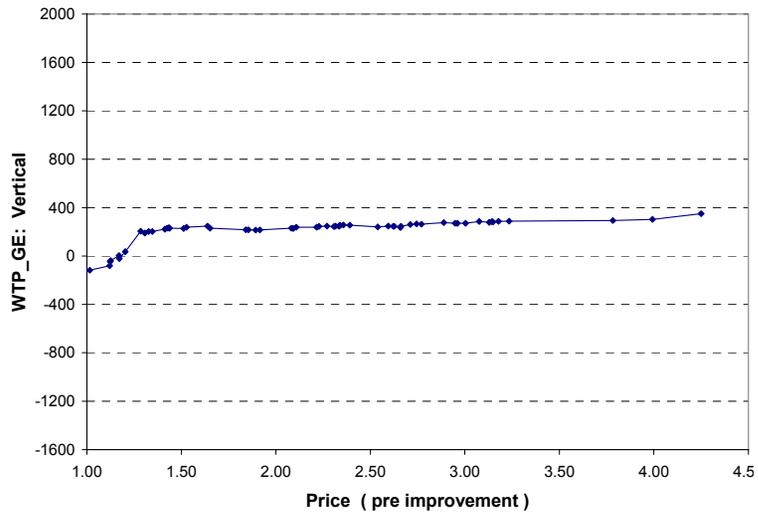
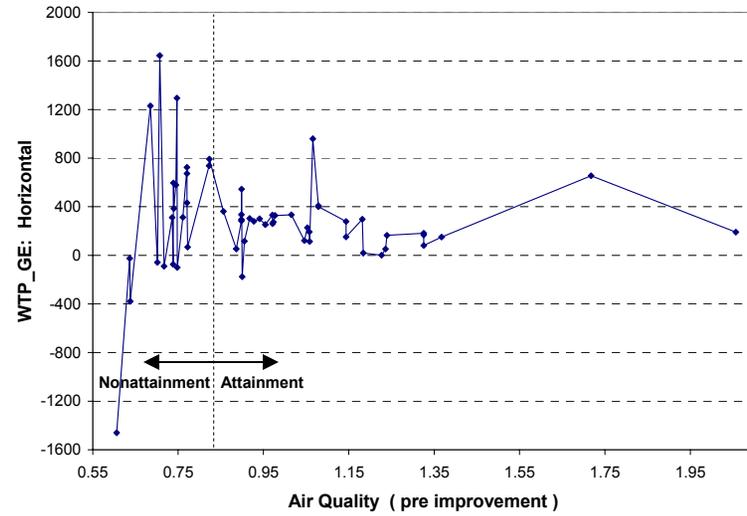
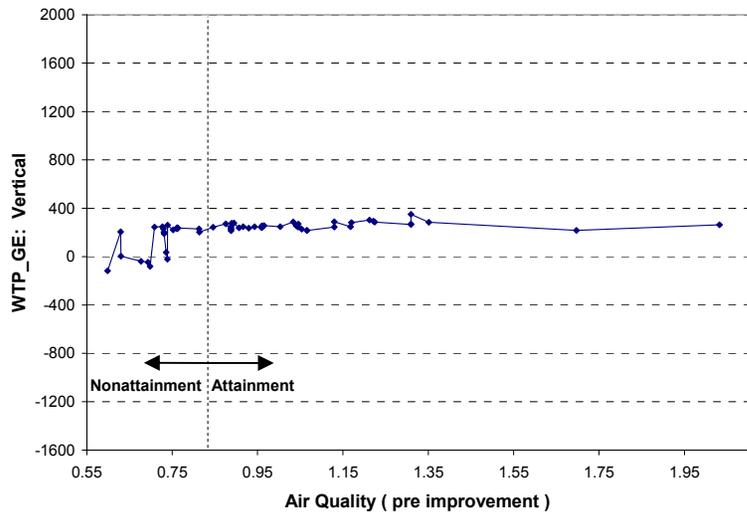


Figure 6.5: Average WTP^{GE} across Communities, by Air Quality, Price, and Depiction of Preference Heterogeneity

Chapter 7: Key Findings and Future Research

I. Summary of Key Findings

50 years ago, Charles Tiebout (1956) compared the choice of a house location to a shopping trip for local public goods. In his own words: “*There is no way in which the consumer can avoid revealing his preferences in a spatial economy. Spatial mobility provides the local public-goods counterpart to the private market’s shopping trip.*” While this influential insight forms the basis for empirical models of sorting behavior, recent evidence provides support for the hypothesis that local public goods are not the dominant factor in determining where households choose to live. If local public goods were the primary factor affecting location choices, a decrease in moving costs would make it easier for households to stratify according to their relative preferences, decreasing the amount of preference heterogeneity within communities. As the diversity in preferences within communities decreases, one would expect them to become more specialized in their provision of local public goods. However, in an empirical analysis of communities in the United States over the past 150 years, Rhode and Strumpf (2003) find just the opposite—as mobility costs have declined over time so has the between-community variation in proxy measures for public good provision such as local taxes per capita and school taxes per capita. Moreover, the annual American Housing Survey consistently reports “proximity to employment” as the reason most frequently cited by households for choosing to live in their current neighborhood.

The previous chapters sought to extend the sorting literature to recognize that a location conveys a set of job opportunities together with local public goods, while simultaneously recognizing that households differ in their job skills and in their relative preferences for those public goods. This new dual-market framework manifests Tiebout’s logic by translating information on households and their location choices in the housing and labor markets into structural parameters that represent their (revealed) preferences for local public goods. Furthermore, the econometric model relaxes the need for a priori assumptions about the shape of distributions for the parameters used to characterize preference heterogeneity in the population of households. The new framework was used to ask the following question: What can we learn from a household’s “public goods shopping trip”?

The key findings can be summarized in terms of four sets of results: conceptual and methodological, econometric, empirical, and policy.

(A) Conceptual and Methodological Results

Given a specification for preferences and a definition for the set of possible locations, a household's observed location provides "set identification" of their preferences. In other words, their choice implicitly defines a set of values for the parameters of the utility function that describe how local public goods contribute to sorting behavior. To attach values from this set to the population of households requires additional assumptions about the distribution of each preference parameter. The indirect utility function can be used together with the data that describe the choice set to partition preference space into regions that define the set of values for the heterogeneous parameters capable of explaining each observed location choice. This process can provide a graphical analog to Samuelson's (1948) revealed preference logic by illustrating how a household's preferences are identified by the choices they make and by the choices they could have made, but did not.

The partition of preference space was used to illustrate how the characterization of the urban landscape and the depiction of preference heterogeneity influence what can be learned about the demand for local public goods. Relaxing vertical differentiation to allow households to differ in their relative preferences for multiple public goods expands the range of consistent explanations for their observed behavior. Recognizing that working households make a joint job-house choice produces a similar effect. While both generalizations can affect conclusions about the demand for local public goods, they also increase the sensitivity of those conclusions to arbitrary assumptions about the shape of the distribution used to describe the heterogeneous preference parameters. This presents the analyst with a type of "bias/variance" tradeoff. More precisely, if the "true" choice problem is characterized by working households with horizontally differentiated preferences making a joint job-house choice, then treating income as exogenous and preferences as vertically differentiated will have two effects. It will lead to biased parameter estimates which have the potential to

produce biased welfare measures and it will decrease the sensitivity of those welfare measures to distributional assumptions.

The characterization of the location choice problem also determines the scope for spatial substitution. In other words, the depiction of preference heterogeneity and the description of the urban landscape both influence empirical predictions for the way a household will adjust its location following a change in public goods. If households are depicted as having vertically differentiated preferences, each community will have exactly two substitutes—the adjacent communities in the ranking by overall public goods provision. The number of substitutes increases when the model is generalized to allow horizontal differentiation and a joint job-house choice. Moving from vertical to horizontal differentiation increases the total number of substitutes by recognizing that individual households may differ in what they perceive as being the closest substitutes for a given community. With horizontal differentiation, the number of substitutes for each community is increasing in the number of heterogeneous attributes that differentiate those communities. Introducing labor market choices increases the number of substitutes because it expands the set of heterogeneous attributes to include wages and commute times.

(B) Econometric Results

Generalizing Epple and Sieg's (1999) model to allow wage income to vary with location and to recognize that households may differ in their relative preferences for multiple public goods required developing a new approach to estimation. The resulting microeconomic framework has three key features. First, it introduces a heterogeneous "job skill" parameter to represent each worker's ability to earn more or less than the average wage paid to other workers in the same occupational category in alternative labor markets. This strategy effectively treats each worker's job opportunity set as a latent variable, recognizing that a worker's job opportunities depend on their (unobserved) idiosyncratic characteristics such as education, experience, and "teamwork skills".

The second key feature of the estimator is that it eliminates the need for a priori assumptions on the distributions for heterogeneous preference parameters. A Gibbs sampling

algorithm was developed to recover an approximation to the partition of preference space. This approach allows the econometrician to distinguish the effects of the identifying power of structural restrictions imposed on the indirect utility function from those of maintained assumptions about the distribution of preferences. Distributional assumptions are still required to translate the information in the partition into a distribution of welfare measures. However, recovering the partition provides a way to analyze the potential sensitivity of the resulting welfare measures to the distributional assumption.

Finally, a general equilibrium framework was developed to simulate how households and markets would adjust to a large-scale change in the provision of public goods. The new locational equilibrium prices, wages, and location choices can be used to calculate general equilibrium measures of the willingness-to-pay for the change. The algorithm can produce divergent solutions and multiple equilibria. Preliminary results suggest that these limitations stem from the maintained assumption that the capital gains from changes in housing prices and wage rates are collected by absentee homeowners and absentee shareholders. Future research could investigate this hypothesis by calibrating the algorithm to reproduce the share of homeowners observed in the population of households. Likewise, the share of households' non-wage income that comes from dividends could be linked to changes in firm profitability that stem from changes in the wage rate.

(C) Empirical Results

The economic consequences of generalizing Epple and Sieg's (1999) framework were evaluated by comparing results from the new dual-market estimator to the results from two special cases—the Epple-Sieg model, and an intermediate “single-market” case that admits horizontal differentiation but treats wage income as exogenous. Each version of the model was estimated using data on households and their location choices in the San Francisco-Sacramento region of Northern California, where communities were depicted as differing in their provision of two observable public goods—air quality and school quality. The results were translated into measures of the average MWTP for an air quality improvement, using ozone concentrations as a proxy measure for air quality. Moving from the Epple-Sieg model

to the dual-market framework increased the average MWTP by 47% under the “naïve” assumption that preferences are distributed uniformly within each region of the partition of preference space. More generally, the range of estimates in the dual-market case (\$33 to \$226) contains the range in the single-market case (\$57 to \$168) which contains the point estimate from the vertical model (\$83). This illustrates the potential economic relevance of the “bias/variance” tradeoff.

Generalizing the characterization of the choice problem had the largest impact on individual communities. For example, consider the Santa Rosa high school district which has above average air quality (71st percentile), below average school quality (40th percentile), and relatively inexpensive housing (28th percentile). In the Epple-Sieg model, the hierarchical ordering of communities by the price of housing implies that households who choose to live in Santa Rosa tend to do so because of its low housing prices. Put differently, the model predicts that households in Santa Rosa have relatively low preferences for public goods relative to the private good component of housing and, on average, a relatively low MWTP for reduced ozone concentrations (\$53). Generalizing the Epple-Sieg model to recognize that households differ in the relative importance they assign to different public goods increases the estimate for the average MWTP to \$114. Intuitively, allowing households to have horizontally differentiated preferences recognizes that the bundle of public goods provided in Santa Rosa will appeal to households who have relatively strong preferences for air quality. These households will also tend to have a relatively high MWTP for reduced ozone concentrations. In addition to clean air, the choice to live in Santa Rosa requires workers in most occupations to accept a lower wage or a longer commute to a different metropolitan area. Therefore, by choosing to live in Santa Rosa despite these drawbacks, working households reveal even stronger preferences for the bundle of public goods it provides. Extending the horizontally differentiated model to depict this joint job-house choice increases the average MWTP to \$216.

(D) *Policy Results*

Measures of the marginal willingness-to-pay for reduced ozone concentrations are not directly comparable with estimates of the MWTP for air quality reported in much of the existing literature where air quality is typically measured by particulate matter or by the number of days during a year that ozone levels exceed state or federal standards. However, assuming that all of these measures are simply different proxies for clean air, they can be compared in terms of a common proportionate change. Converting the range of values from the existing literature into measures that would be equivalent to a marginal (1 part-per-billion) reduction in ambient ozone concentrations implies a range from \$7 to \$154. Under the uniform distribution assumption, the corresponding estimate from the single-market version of the model was \$109, compared to \$122 in the dual-market case. These results are based on estimates for the price and income elasticities of housing that are consistent with benchmark ranges. Nonetheless, it is important to acknowledge that the maintained assumptions of the sorting framework differ from the alternative assumptions that underlie the (reduced-form) hedonic models which provide bounds on the range of existing estimates.

The estimation results were used to analyze the welfare implications for a hypothetical air quality improvement that would allow the entire San Francisco-Sacramento region to meet California's new 0.07 ppm standard for ozone concentrations. First, the uniform distribution of preferences recovered during the estimation was used to calculate each household's ex-ante willingness-to-pay for the improvement. Then the general equilibrium framework was used to simulate how households and markets would adjust to the change. Finally, the new equilibrium housing prices, wage rates, and location choices were used together with the new ozone levels to calculate a theoretically consistent measure for each household's ex-post willingness-to-pay for the (non-marginal) air quality improvement. These welfare measures differ from the results that would be obtained by following the standard practice of multiplying an estimate for the average MWTP by the size of the improvement. The importance of this difference can be illustrated by converting the ex-ante and ex-post welfare measures into unit values that are comparable to measures for the MWTP. In the single-market case, dividing the aggregate (ex-ante) willingness-to-pay by the

average ozone reduction times the number of households yields a unit value benefit of \$36 per household per 1 ppb reduction—approximately 1/3 the average MWTP (\$109). Thus, multiplying the average MWTP by the size of the ozone reduction would overstate the ex-ante willingness-to-pay by approximately 200%. The (ex-ante) unit value benefit increases to \$41 in the dual-market case, which is also approximately 1/3 the average MWTP (\$122). Accounting for the welfare implications of general equilibrium adjustment increases the unit value to \$91. While the cycle of price and wage increases that appears to drive this measure may cause it to be an overestimate, it still falls below the average MWTP.

The difference between the unit value benefit measures and the average MWTP highlights the economic importance of the sorting logic that underlies the estimation. This logic implies that, all else held constant, households with the weakest preferences for air quality and the lowest willingness-to-pay for an improvement will choose to live in the communities with the highest ozone concentrations. Since these communities are the ones that have the largest air quality improvements in the hypothetical attainment scenario, the aggregate benefits from the ozone reduction are much smaller than if air quality were improved uniformly within the study region. The opposite would be true if the largest ozone reductions were to occur in the communities with the highest air quality.

II. Future Research on the Location Choice Problem

While the idea of using households' location choices to reveal their demand for spatially differentiated environmental amenities dates back at least Ridker's (1967) air pollution study, structural sorting models appear to be at an early stage of development. Parts (A), (B) and (C) of this section suggest ways to improve future applications to the housing and labor markets by further exploiting the information contained in the composite unobserved public good recovered for each community, obtaining micro data on individual households, and developing new strategies to bound the partition of preference space. Parts (D) and (E) discuss how the estimator and the general equilibrium framework could be extended to

analyze policy issues for a larger geographic area and to investigate the performance of alternative, reduced-form approaches to estimation.

(A) *Reinterpreting the Index of Unobserved Public Goods ($\bar{\xi}_j$)*

Recall that the set of estimates for $\bar{\xi}$ “pick up the slack” in explaining observed location choices during the estimation. The empirical importance of this term was underscored by the result that households’ location choices could not be explained on the basis of observed public goods for 88 of the 122 communities in the San Francisco-Sacramento region. While this statistic has not been reported in previous sorting applications, similar results have been found in other applications of structural discrete-choice models (e.g. Bajari and Benkard [2005]). Given the apparent influence of unobserved attributes on observed choices, future research could reconsider the interpretation given to the $\bar{\xi}_j$ index and investigate how its constituent elements influence what is learned about the demand for the observed public goods.

Throughout the previous chapters, $\bar{\xi}_j$ was interpreted as the composite effect from a bundle of public goods that households observe but the analyst does not. However, even within a relatively small geographic area, households’ location choices may also depend on “unobserved private goods” such as proximity to family, friends, and a familiar environment. Likewise, some households may prefer to live in communities where a majority of the other households share a similar education, income, or ethnic background (Bayer, McMillan, and Reuben [2005]). Thus, unobserved public goods, unobserved private goods, and the distribution of demographic characteristics for the population of households may all contribute to the set of estimates for $\bar{\xi}$. The relative contributions of each could be analyzed by regressing the estimates for $\bar{\xi}$ on a set of community-specific proxy measures for the (formerly) unobserved public goods and demographics. Equation (7.1) depicts a linear version of this regression, where *public* and *demographic* represent vectors of public and demographic characteristics in community j , and the residual (ψ_j) includes the composite

effect of unobserved private goods.

$$(7.1) \quad \bar{\xi}_j = \omega_1 \cdot public_j + \omega_2 \cdot demographic_j + \psi_j.$$

For example, in the application to Northern California it was hypothesized that the set of estimates for $\bar{\xi}$ reflect variation in climate, access to cultural amenities, and proximity to parks. Measures for each could be included as right-hand side variables in the reduced-form regression together with summary measures describing the demographics for the population of households in each community (e.g. average income, average age, population share by race).

In addition to providing a diagnostic check on the model, the regression results may suggest areas where it could be revised. Recall that households are restricted to have vertically differentiated preferences for the set of community-specific attributes measured by $\bar{\xi}$. Thus, the public goods index can be divided into a set of “horizontal” characteristics (air quality and school quality) and a set of “vertical” characteristics (climate, cultural amenities, parks, demographics, and ψ). If demographics are found to explain much of the variation in $\bar{\xi}$, it may make sense to redefine them as horizontal characteristics in order to recognize that households differ in how they evaluate the attributes of their neighbors. Of course, this modeling choice would also present the analyst with a new dimension of the bias/variance tradeoff described earlier. Alternatively, if public goods and demographic characteristics are found to explain relatively little variation in $\bar{\xi}$, future research might consider revising the model to explicitly acknowledge the role of unobserved private goods in households’ location choices.

The regression results could also be integrated into the general equilibrium framework. For example, if households react to a change in the provision of a public good by moving, the distribution of demographic characteristics within each community may also change, requiring an adjustment to the public goods index. Introducing this “feedback” loop into the simulation process would require additional data on the demographic characteristics

of individual households, conditional on their income and occupation. While these data are not publicly available, they are reported in the restricted Census micro data.

(B) Using Census Micro Data to Describe Households and their Choices

Publicly available Census data were used to classify the 3.2 million households in the San Francisco-Sacramento region according to their locations, non-wage incomes, and the occupational categories of their primary earners. However, this classification process required numerous assumptions about the properties of the distributions of wage and non-wage income. Furthermore, the resulting data imposed four constraints on the estimation. First, while the demand for housing is presumed to depend on permanent income, the housing expenditure function was estimated using data on current income. Second, while 51% of working households in the region have more than one wage earner, the wage income from “secondary earners” was included in exogenous non-wage income. Third, each worker’s wage-commute options were defined on the basis of their occupational category rather than their specific occupation. Finally, the estimation relied on a discrete approximation to the distribution of income in the population of households. Each of these limitations could be addressed in future applications by obtaining access to the restricted Census micro data available only at Census data centers.

Although the Census micro data do not contain a direct measure of wealth or permanent income, they include detailed information on each household’s sources of income (e.g. wages, dividends, rental income). In addition, they contain detailed demographic information on each household, such as the ages and levels of education of each of its workers. This information could be used to construct an estimate for each household’s permanent income, following the strategy used by early studies on the demand for housing (Polinsky [1977]). Developing a measure for permanent income would improve the estimation of the homogeneous housing demand parameters in two ways. First, it would improve the efficiency of the estimation by allowing retired households to be included. Second, it would eliminate any bias that stems from using current income as a proxy for permanent income for working households. Of course, these measures for permanent income

will not necessarily reflect a household's wealth portfolio. That is, a household may choose to invest in property or other assets that do not contribute to current income but influence the demand for housing. Future research could begin to address the implications of this potential omission by investigating the relationship between the demand for housing, wealth, and permanent income.

More detailed information on workers' occupations would help to define the set of wage-commute options faced by each working household. For example, rather than classifying a worker as having an occupation in the "architecture and engineering" category, they could be classified according to their specific job title such as "surveying and mapping technician" or "civil engineer". This would increase the number of worker "types" from 22 to 821. Wage data for each of these 821 occupations are available by PMSA. By providing a more accurate depiction of each worker's wage-commute options, the data on individual job titles could reduce the reliance on unobserved job skill to help explain observed choices⁸⁵. The Census micro data would also allow the empirical model to be extended to consider a dual-earner job search. That is, for households with two workers, the set of wage-commute options could be defined to recognize that wages for each depend on their job locations. However, this extension would also raise new modeling issues such as whether each worker should be depicted as having a separate job skill parameter and a separate opportunity cost-of-time. Similarly, one might wish to reconsider the objective function used to define sorting behavior. For example, do dual-earner households make location choices to maximize total household income, to maximize the wage income of the primary earner, or something in between?

Moving from aggregate data on the distribution of households to micro data describing those households significantly improved parameter estimates in past microeconomic applications (e.g. Petrin [2002], Berry, Levinsohn, and Pakes [2004]).

⁸⁵ Although the Census does not report data on the attributes of individual jobs (e.g. health coverage, hours worked, risk of injury/death) if data on these attributes were available from another source they could be used together with the Census data on wage income to adjust the wage rates that characterize the set of job choice alternatives. The idea would be to calculate a set of fixed effects that represent the "effective" wage rate in each (occupation, labor market) combination. This calculation would parallel the process used to recover the price of a homogeneous unit of housing in each community. Access to micro data on wages and job attributes would also create more flexibility in the way that a "labor market" is defined.

However, in the present model, it would also complicate the estimation process. With all the possible (occupation, non-wage income, location) combinations in the Census micro data, no two households may be alike. Thus, each individual household could have a distinct region in the partition of preference space. Developing a strategy to characterize 3.2 million regions of preference space would raise a new set of computational challenges. One strategy to simplify the problem would be to treat workers in each occupation as having homogeneous job skills. In this case, the set of job opportunities faced by workers in each occupation could be estimated from a hedonic wage equation, using the Census micro data. This would require estimating a reduced-form expression that explains how wages vary with the attributes of workers (e.g. race, sex, age, education, occupation) and the attributes of their jobs (e.g. non-wage benefits, mortality risks, morbidity risks). The resulting expression could then be used to predict the wage each working household would be paid in each alternative labor market, as in Bayer, Keohane, and Timmins (2006). However, an important limitation of this approach is that there is no guarantee the predicted wage rates would be capable of rationalizing observed job locations unless idiosyncratic job-specific “taste” parameters were also added to the utility function.

(C) Bounding the Partition of Preference Space

Recovering an approximation to the partition of preference space provides a new perspective on measures of the willingness-to-pay for improvements in public goods by allowing the analyst to investigate the sensitivity of welfare measures to assumptions about the distributions of heterogeneous preference parameters in the population of households. However, communities with extreme provision of one or more public goods tend to have unbounded regions in the partition. This was addressed during the estimation by placing an (arbitrary) \$500 upper bound on the ex ante willingness-to-pay for a 1 part per billion reduction in ozone concentrations. However, the choice for the upper bound can have a large impact on welfare measures for households who live in the corresponding communities. Thus, future research should develop strategies for using more information to provide bounds on the partition. One possibility would be to conduct a stated preference study in the

communities with unbounded regions. However, this could pose challenges for reconciling the stated and revealed preference data. Two other possibilities include: (a) exploiting weak complementarity together with a restriction on the minimum house size, and (b) defining a set of “outside locations”.

The revealed preference logic that underlies the estimation implicitly treats the private good component of housing as a *weak complement* for public goods (Mäler [1974]). In other words, for a household to obtain utility from the public goods provided by community j , the household must live there. To see how combining weak complementarity with a restriction on the minimum house size can bound the partition, suppose the region of preference space that corresponds to community j is unbounded in the α dimension. Therefore, as α_i increases so does the value for p_j that would make household i exactly indifferent between community j and their second favorite community. Let h^* represent the minimum quantity of housing that can be occupied in any community and let $p_{i,j}^*$ represent the price of housing in community j that would induce a household with income of y_i to consume exactly h^* units of housing. In other words $p_{i,j}^*$ is the “choke price”, above which the household cannot afford to live in community j .

Weak complementarity requires the level of utility from the (hypothetical) situation where household i consumes fewer than h^* units of housing in community j to be less than the utility from consuming public goods at observed prices in every other community, as depicted in equation (7.2).

$$(7.2) \quad V(p_{i,j}^* + \varepsilon, g_j | y_i, \alpha_i, \gamma_i) < V(p_q, g_q | y_i, \alpha_i, \gamma_i) \quad \forall q \quad \text{and} \quad \varepsilon > 0 .$$

Intuitively, the expression requires households to prefer occupying a house in a community with low provision of public goods to being homeless in a community with high provision of public goods. Given the normalization that the weights in the public goods index sum to 1, the system of equations in (7.2) can be solved to provide an upper bound on the theoretically consistent range of values for α .

While the bounds derived from this strategy would be theoretically consistent, they would have two limitations. First, h^* represents a specific value for the index of housing and, therefore, does not have a direct economic interpretation. This could be addressed by using the first-stage estimates for the housing demand function to calculate the aggregate quantity of housing in the study region. Dividing by the total square footage of housing would provide of measure for h^* per square foot. Multiplying the resulting figure by the minimum number of square feet (e.g. 400) would produce a value for h^* that could be used to calculate a bound on α . Nevertheless, these bounds may correspond to partial equilibrium welfare measures that exceed income.

The second strategy for bounding the partition would expand the choice set to include a set of “outside communities” that provide more public goods than the communities in the study region. For example, a household with extremely strong preferences for air quality might decide to live in Lake County rather than the San Francisco-Sacramento region. Located 140 miles northeast of San Francisco, Lake County has not exceeded state standards for daily ozone concentrations during the past 20 years. In comparison, the San Francisco air basin averaged 26 violations a year between 1985 and 2005, while the Sacramento air basin averaged 59 violations. Similarly, a household with extremely strong preferences for school quality may decide to move to the San Marino Unified school district in Los Angeles, which reports the highest value for California’s Academic Performance Index.

Including Lake County and San Marino in the choice set would violate the maintained assumption that households cannot move outside the San Francisco-Sacramento region. One way to address this problem would be to incorporate moving costs. This strategy can be illustrated by returning to the 4-community example that was used throughout the earlier chapters. Table 7.1 displays the price of housing and provision of public goods in each of the four communities. Recall that when the choice set is defined as $\{A, B, C, D\}$, the region of preference space that corresponds to community C is unbounded, as depicted in panel A of figure 7.1. Let E represent a community outside the study region that provides higher values of both public goods. Suppose the annualized cost of moving to region E from any of the other four communities is \$1700. This could represent a combination of lower

wage income and the physical, information, and psychic costs of moving. Panel B in figure 7.1 illustrates how, given this moving cost, including community E in the choice set bounds region C of the partition.

A theoretically consistent approach to using outside communities to bound the partition of preference space would also need to address the geographic implications of extending the choice set. For example, if households can move from San Francisco to San Marino, they should also be capable of moving to San Luis Obispo, which is located halfway in between. This could be addressed by compiling data on the contiguous region from San Marino through Lake County, and using it to define two sets of communities: “inside communities” and “outside communities.” The model would be estimated for households living in the inside communities, using the choice set defined by the union of inside and outside communities. As in the current model, the cost of moving between inside communities could be treated as negligible. The cost of moving from inside communities to outside communities could be estimated using specifications from the existing literature on migration (e.g. Herzog Jr. and Schlottmann [1981]).

(D) Moving Costs and Expanding the Geographic Scope of the Study Region

Including moving costs in the budget constraint would also allow the model to be extended to larger geographic areas. This could provide more variation in the choice set and expand the scope for policy analysis. For example, if housing communities were defined as counties, the model could be estimated from data on one or more states, providing a structural analog to the interregional hedonic literature⁸⁶. Recent evidence has suggested that accounting for moving costs in a national analysis can have a large impact on welfare measures. For example, Bayer, Keohane, and Timmins (2006) report that including moving costs in their version of the interregional hedonic model triples their estimate of the average MWTP for air quality.

⁸⁶ Given the cost and difficulty of collecting micro data on individual housing transactions across a large region of the country, the first-stage of the estimation would have to be adapted to use a different source of data on housing prices and their structural characteristics. One possibility would be to use the limited information on housing characteristics contained in the Census PUMA data together with the values that homeowners report for their homes.

If moving costs prevent a large share of the population from locating in the region of the country that would be their preferred choice were they freely mobile, the resulting locational equilibrium will be hysteretic. Graphically, hysteresis would imply that the borders of the partition overlap. Figure 7.2A illustrates this using the four-community example where the four communities are now interpreted as counties in different parts of the country and the annualized cost of moving between counties is set to \$350. Regions of the partition denoted by the bold letters A, B, C, and D have the same interpretation as in earlier partitions. For example, any household with preferences in region A will maximize its utility by locating in community A. Households with preferences in the 7 regions of the partition that have labels beginning with the letter A (other than region A), such as AC, ACD, and ADB, would choose to live in community A if moving costs were zero. However, if they were born in the communities denoted by the other letters in the labels, the annualized cost of moving would prevent them from migrating to A. Likewise, some households living in community A may have preferences in regions of the partition like DBA, indicating that they prefer the combination of public goods and housing prices provided by communities D and B but are not willing to pay the cost of relocating there.

During the estimation, the partitioning process could be performed separately for recent movers and longtime residents. For recent movers, the Gibbs sampling algorithm would be implemented just as before, sampling over each of the regions in figure 7.2B. For households who have not moved recently, the annualized cost of moving expands the area of preference space that corresponds to each observed location choice. For example, for households living in community D, the Gibbs algorithm would sample over the area defined by the union of all 19 regions in figure 7.2A that contain the letter D in their labels.

Accounting for moving costs during the estimation would increase the sensitivity of welfare measures to distributional assumptions by expanding the size of the preference regions capable of explaining each observed choice. This effect would be analogous to the way that including job opportunities in the model was found to “stretch” the preference regions. In other words, like job opportunities, moving costs increase the range of consistent explanations for observed location choices. If this logic were to be implemented in a national

analysis of the demand for improved air quality, the results would provide a different perspective on the consequences of introducing moving costs than the analysis by Bayer et al. One interpretation would be that moving costs decrease the precision with which we are able to estimate the demand for air quality (or other amenities) from national data. Another interpretation would be that moving costs limit or prevent spatial integration between local markets for air quality in different parts of the country and therefore, there is no national market for air quality.

One could also argue that mobility tends to be highly correlated with occupation. For example, when a worker in the *managerial* or *architecture and engineering* occupational categories moves to a new job in a different part of the country, their new employer may pay for their physical moving costs. This is less likely to be true for workers with jobs in the *construction* and *production* categories. As a result, it may make sense to think of there as being a national market for labor in some occupations, but not others. This could be addressed by using data on national migration patterns by occupation to define the geographic scope of the choice set separately for each occupation, following the logic that was used to define the study region in chapter 4.

(E) *Preference Heterogeneity, Capitalization, and Reduced-Form Estimation*

The general equilibrium framework could provide another perspective on the interregional hedonic literature by forming the basis for a simulation-based study to investigate how heterogeneity in preferences and skills affects the ability of reduced-form estimators to produce consistent estimates of the average MWTP for improvements to public goods. This analysis would parallel Cropper, Deck, and McConnell's (1988) study on the performance of different specifications for the hedonic price function. Following their logic, the general equilibrium model could be used to simulate shocks to the provision of public goods and then solve for a new locational equilibrium. The new equilibrium housing prices and wage rates would then be treated as "data" and used together with the simulated data on each public good to estimate the marginal implicit prices for those public goods. The difference between the marginal implicit prices estimated from the reduced-form model and the average MWTP

in the population of households would provide a measure of performance for the reduced-form estimator. This exercise could be repeated under different assumptions about the nature of preference heterogeneity, job skills, and moving costs.

A simulation-based analysis could also provide a new perspective on the performance of the traditional interregional hedonic model relative to the recent “difference-in-difference” applications. Recall that the intuition for the interregional hedonic model is that households must “pay” for improvements to public goods through higher rents, lower wages, or both (Rosen [1979]). Applications of the model using cross-section data have found that both rents and wages reflect a substantial share of the total implicit price of public goods (Blomquist et al. [1988] and Graves and Waldman [1992]). That is, they find a negative relationship between wage rates and the level of public goods provided (including measures of air quality). More recently, interregional hedonic applications that have used time-series data and a “difference-in-difference” approach to estimation have found that housing prices tend to increase following air quality improvements, but wage rates do not. Chay and Greenstone (2005) find that air quality improvements have no effect on wage rates while Bayer et al. (2006) find a positive relationship. The simulation framework could be used to investigate whether these seemingly conflicting results can be explained by heterogeneity in preferences and job skills. More generally, the general equilibrium model could be used to investigate the relative performance of the two reduced-form approaches to estimation under specific assumptions about preferences, job skill, the supply of housing, the demand for labor, and moving costs.

III. Broader Implications for Future Research

Some of the key results from this research extend beyond the location choice problem. In particular, the conceptual and methodological results apply generally to structural discrete-choice models of the demand for a differentiated product. As such, they have implications for the recreation demand literature which depicts individuals “sorting” across a set of recreation sites. Similarly, some of the lessons learned from modeling sorting behavior and a

multidimensional equilibrium may be useful in other problems within public and environmental economics. The remainder of this section considers a few possibilities.

(A) *Structural Discrete-Choice Models / Recreation Demand*

The sorting and recreation demand literatures depict parallel choice problems. In the recreation choice problem, households that differ in their preferences and wage income choose from a discrete set of recreation sites that differ in a set of characteristics describing opportunities for recreation and amenities that affect the quality of recreation. As in the dual-market location problem, the price structure for the set of choice alternatives is effectively a latent variable because the cost of traveling to a recreation site depends on the agent's (unobserved) opportunity cost of time. Given these sources of heterogeneity, recreation choices are typically modeled within a random utility framework, using a priori shape assumptions to characterize the distribution of preference parameters used to represent unobserved heterogeneity. Recent work has extended this framework to recognize that the analyst may be unable to observe some site characteristics (Murdock [2006]).

Given the similarities between the two problems, many of the methodological and econometric results from the sorting literature should be applicable to recreation demand problems. For example, the partitioning process outlined in chapter 3 could be extended to relax the need for a priori distributional assumptions in a recreation demand framework. That is, given a form for the indirect utility function, a Gibbs sampling algorithm could be developed to recover an approximation to the partition of preference space. The resulting partition could then be used to test the sensitivity of the resulting welfare measures to distributional assumptions. Since recreation demand studies are often based on a small number of choice alternatives, it seems probable that distributional assumptions will have a significant effect on welfare implications⁸⁷.

The data sets used in recreation demand studies also present an opportunity to evaluate the “out-of-sample” predictive power of models hypothesized to represent consumer

⁸⁷ For example, half the studies summarized in Phaneuf and Smith's (2005) discussion of recreation demand models have fewer than 70 recreation sites (see their table 3.1).

behavior. For example, Provencher and Bishop (2004) use repeat observations on anglers who make trips to Lake Michigan to compare the predictive power of a random utility model under different depictions of preference heterogeneity. The same logic could be used to evaluate the predictive power of a random utility model under different distributional assumptions for the heterogeneous parameters *conditional* on a depiction of preference heterogeneity (i.e. vertical or horizontal). Similarly, natural experiments such as hurricanes, fires, or floods that change the set of recreation opportunities present an opportunity to observe how the depiction of preference heterogeneity and distributional assumptions affect the model's ability to predict how individuals adjust to the shock. Future research could consider whether the lessons learned from this type of study could also be used to inform the way the sorting literature depicts household behavior.

(B) *Multidimensional Equilibria*

There are many areas in public and environmental economics where individuals have the ability to adjust their behavior along multiple dimensions following a change in the quality of public goods. For example, individuals may react to a change in the quality of recreation attributes at a particular site by changing the length of their visit or by choosing to visit a different site. Similarly, homeowners may respond to an increase in the perceived risk of natural disasters by purchasing additional insurance or by taking actions to fortify their home. Like the dual-market locational equilibrium concept, these examples can be envisioned as multidimensional equilibria. In the recreation example, the consumer's decision can be envisioned as a (site, visit duration) pair. In the natural disaster example, a choice can be defined as a unique combination of expenditures on (insurance, fortification). The results from the general equilibrium simulations in chapter 6 suggest two lessons for modeling shocks to multidimensional equilibria when consumers have multiple opportunities for adjustment.

First, the depiction of preference heterogeneity can influence the diversity of adjustment paths. Increasing the number of dimensions in which households differ tends to increase the number of substitution possibilities associated with each choice in the model.

Second, there may be feedback effects between adjustments made along different dimensions of the equilibrium. This was illustrated by the cycle of price and wage increases that followed the ozone reduction during the dual-market version of the simulation. In the context of recreation demand, changes in the number of visitors to a particular site (i.e. congestion) may cause people to change the site they visit (as considered by Timmins and Murdock [2005]) or it could lead them to adjust the amount of time they spend at that site. In the natural disaster case, an increase in insurance expenditures in disaster-prone areas could increase insurance premiums elsewhere, leading to an increase in expenditures on fortification. Finally, these two lessons are interrelated. Increasing the degree of preference heterogeneity increases the scope for feedback effects by increasing the diversity of responses to substitution possibilities.

(C) *Heterogeneity and Sorting by Firms*

Heterogeneous firms employ diverse technologies that differ in the environmental externalities they produce. To some extent, variation in technology reflects variation in regulation. However, it may also reflect variation in skill, preferences, and attitudes toward risk. One can envision firms “sorting” into technologies according to these sources of (unobserved) heterogeneity. In this context, the conceptual and methodological results on the depiction of preference heterogeneity in the location choice problem could have implications for regulating non-point source pollution that arises as a byproduct of the production process.

Agriculture provides an example of an industry where heterogeneity in firms is likely to be important. For many crops, farmers choose from a diverse set of production technologies, and past research has modeled this choice process in a random utility setting (e.g. Klonsky and Smith [2002]). Generalizing past work to consider more flexible depictions of heterogeneity and relaxing a priori distributional assumptions could provide a new perspective on the welfare implications of regulations that aim to reduce non-point source pollution by restricting the production process. For example, strawberry farmers in California’s Monterey and Santa Cruz counties choose a bundle of methods for pest and disease control. The method they select will affect the (per/unit) price of their crop, their

yield, and the surrounding environment. If they use methyl bromide as a fumigant, they are likely to obtain a higher yield and a lower price for their crop compared to organic production methods (Bolda et al. [2003, 2004]). However, methyl bromide is being phased out of production because it depletes the ozone layer. A “sorting” framework could be developed to identify a set of values for strawberry farmers’ heterogeneous skills, preferences, and risk attitudes that would rationalize their chosen bundle of production methods. The substitution patterns implied by this partition could be used to predict how farmers would adjust their production practices in response to the phase-out of methyl bromide. Finally, the predicted change in production methods could be used to analyze the welfare implications to farmers as well as to consider the environmental implications of the alternative production methods they adopt.

IV. Conclusions

This chapter summarized key findings from the dissertation and discussed topics for future research. While moving to a dual-market framework provided more information to help identify households’ preferences for local public goods, it also increased the range of consistent explanations for their observed behavior. More precisely, allowing households to have horizontally differentiated preferences and recognizing that job opportunities influence sorting behavior both increased the potential diversity in relative preferences for public goods within each community. This diversity can help to explain Rhode and Strumpf’s (2003) observation that communities have become more alike over time, despite the apparent decrease in moving costs.

Future research on the location choice problem should consider how additional information can be used to identify households’ preferences more precisely. This could be done using restricted Census micro data on the characteristics of individual households and their choices or by developing strategies to provide “natural” bounds on the partition of preference space. Another direction for future research on the location choice problem would be to use the general equilibrium model to evaluate the performance of reduced-form

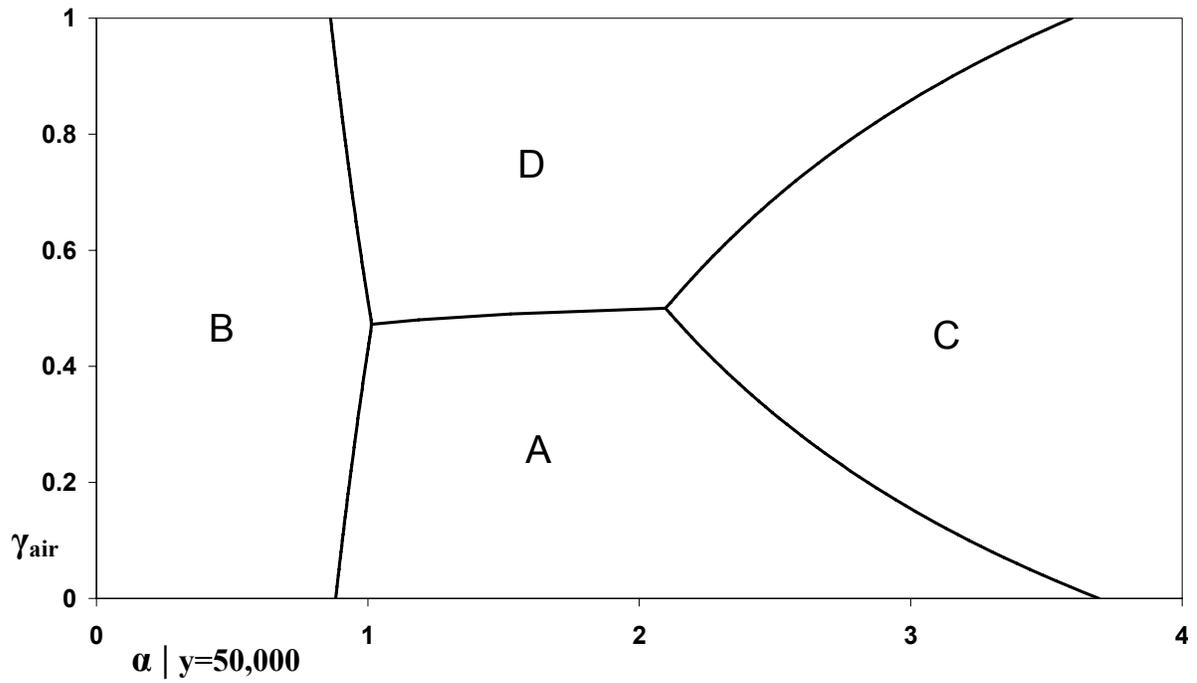
estimators. More generally, the ability to distinguish the extent to which partial equilibrium welfare measures are identified by preference restrictions relative to distributional assumptions may be useful in other economic contexts where structural discrete-choice models are used. Similarly, the characterization of “sorting” behavior and the properties of a multidimensional equilibrium may be broadly applicable in characterizing how non-market equilibria arise in a wide range of other mixed discrete/continuous choice problems.

Table 7.1: Four Community Example Revisited

Community	Public Goods*		Price
	Air quality	School quality	
A	1.25	1.25	1.00
D	1.85	1.65	1.25
B	1.66	1.86	1.26
C	2.00	2.00	1.50
E	2.10	2.10	1.30

* Higher values indicate higher quality.

A. Partition from 4-Community Example, with Region C Unbounded



B. Using Community E to Bound Region C of the Partition (\$1700 moving cost)

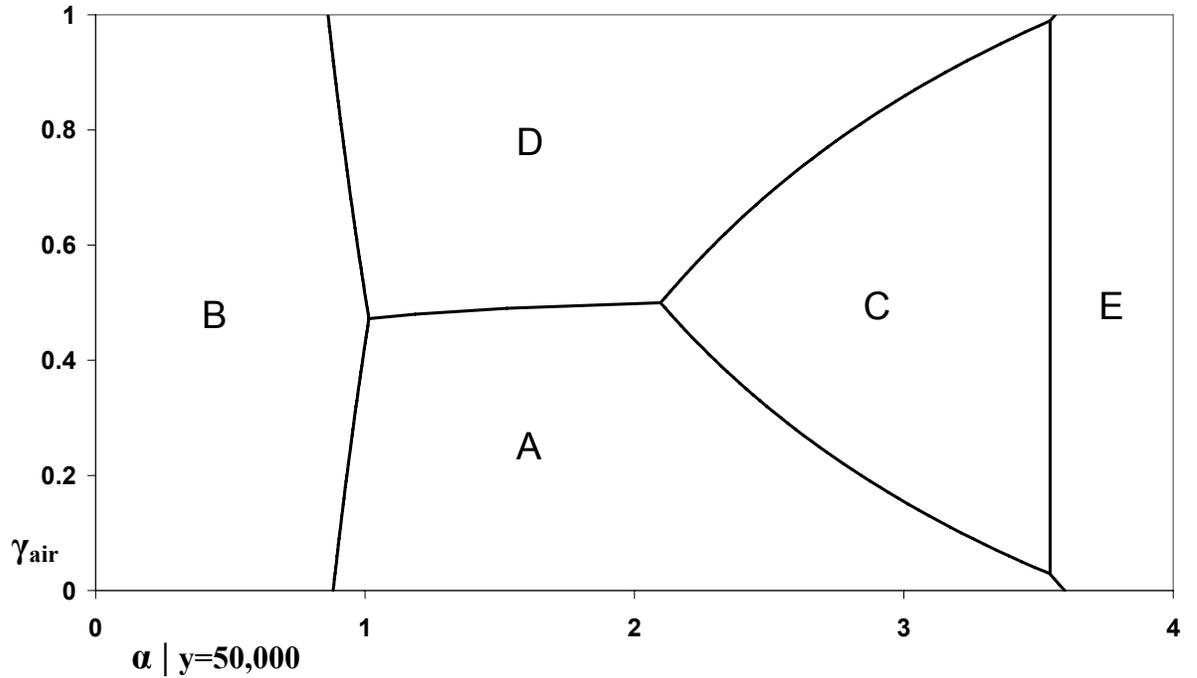


Figure 7.1: Using “Outside Communities” to Bound the Partition of Preference Space

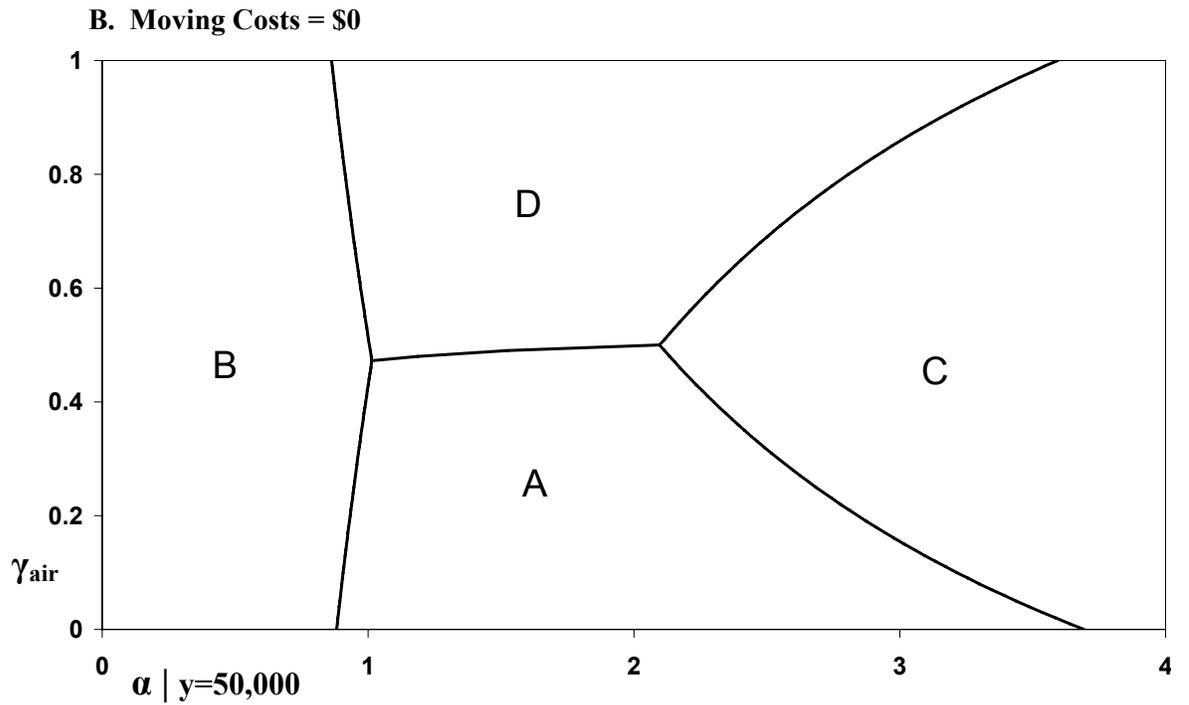
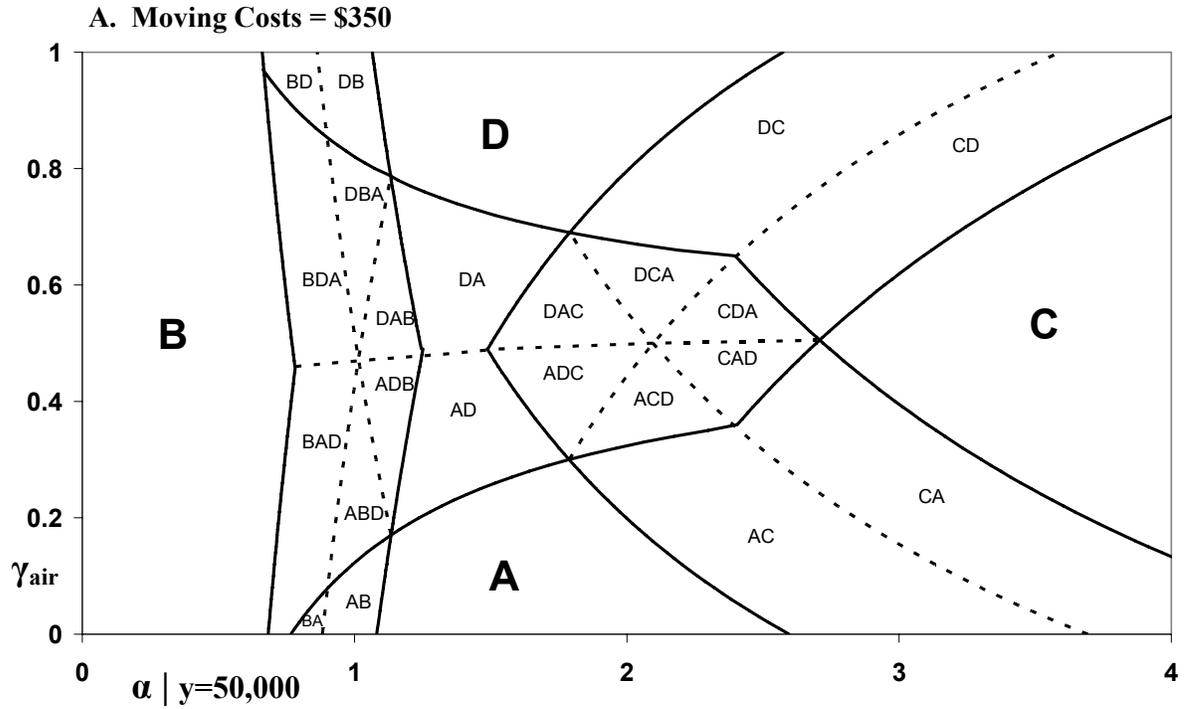


Figure 7.2: Partition of Preference Space with and without Moving Costs

References

I. Journal Articles, Books, and Working Papers

Anderson, Simon P., André de Palma, and Jaques-François Thisse. *Discrete Choice Theory of Product Differentiation*. The MIT Press. 1992

Bajari, Patrick and C. Lanier Benkard. "Demand Estimation with Heterogeneous Consumers and Unobserved Product Characteristics: A Hedonic Approach." *NBER Working Paper #10278*, January, 2004.

_____. "Demand Estimation with Heterogeneous Consumers and Unobserved Product Characteristics: A Hedonic Approach." *Journal of Political Economy*, Vol. 113. (December, 2005): 1239-1276.

Bajari, Patrick and Matthew E. Kahn. "Estimating Housing Demand with an Application to Explaining Racial Segregation in Cities." *Journal of Business and Economic Statistics*. Vol. 23 (January, 2005): 20-33.

Banzhaf, Spencer H. and V. Kerry Smith. "Meta Analysis in Model Implementation: Choice Sets and the Valuation of Air Quality Improvements." Forthcoming in the *Journal of Applied Econometrics*.

Bartik, Timothy J. "The Estimation of Demand Parameters in Hedonic Price Models." *Journal of Political Economy*. Vol. 95. (February, 1987): 81-88.

_____. "Measuring the Benefits of Amenity Improvements in Hedonic Price Models." *Land Economics*. Vol. 64. (May, 1988): 172-183.

Bayer, Patrick, Robert McMillan, and Kim Reuben. "An Equilibrium Model of Sorting in an Urban Housing Market." *Working Paper*. June, 2005.

Bayer, Patrick and Christopher Timmins. "On the Equilibrium Properties of Locational Sorting Models." *Journal of Urban Economics*. Vol. 57. (May, 2005): 462-477.

Bayer, Patrick, Nathaniel Keohane, and Christopher Timmins. "Migration and Hedonic Valuation: The Case of Air Quality." *NBER Working Paper Series #12106*. March, 2006.

Berry, Steven, James Levinsohn, and Ariel Pakes. "Automobile Prices in Market Equilibrium." *Econometrica*. Vol. 63 (1995): 841-890.

_____. "Differentiated Products Demand Systems from a Combination of Micro and Macro Data: the New Car Market." *Journal of Political Economy*. Vol. 112 (January, 2004): 68-105.

- Blomquist, Glenn C., Mark C. Berger and John P. Hoehn. "New Estimates of Quality of Life in Urban Areas." *American Economic Review*. Vol. 78 (March, 1988): 89-107.
- Brown, James N. and Harvey S. Rosen. "On the Estimation of Structural Hedonic Price Models." *Econometrica*. Vol. 50. (May, 1982): 765-768.
- Chay, Kenneth Y. and Michael Greenstone. "Does Air Quality Matter? Evidence from the Housing Market." *Journal of Political Economy*. Vol. 113 (March, 2005): 376-424.
- Chattopadhyay, Sudip. "Estimating the Demand for Air Quality: New Evidence Based on the Chicago Housing Market." *Land Economics*. Vol. 75. (February, 1999): 22-38.
- Cropper, Maureen L., Leland B. Deck, and Kenneth E. McConnell. "On the Choice of Functional Form for Hedonic Price Functions." *Review of Economics and Statistics*. Vol. 70. (November, 1988): 668-675.
- Ekeland, Ivar, James J. Heckman, and Lars Nesheim. "Identification and Estimation of Hedonic Models." *Journal of Political Economy*. Vol. 112 (February, 2004): S60-S109.
- Ellickson, Bryan. "Jurisdictional Fragmentation and Residential Choice." *American Economic Review*. Vol. 61. (May, 1971): 334-339.
- Epple, Dennis. "Hedonic Prices and Implicit Markets: Estimating Demand and Supply Functions for Differentiated Products." *Journal of Political Economy*. Vol. 95 (February, 1987): 59-80.
- Epple, Dennis, Radu Filimon, and Thomas Romer. "Equilibrium Among Local Jurisdictions: Toward an Integrated Treatment of Voting and Residential Choice." *Journal of Public Economic*. Vol. 24 (1984): 281-308.
- Epple, Dennis, Radu Filimon, and Thomas Romer. "Existence of Voting and Housing Equilibria in a System of Communities with Property Taxes." *Regional Science and Urban Economics*. Vol. 23 (1993): 585-610.
- Epple, Dennis, Michael Peress and Holger Sieg. "Identification and Semiparametric Estimation of Equilibrium Models of Local Jurisdiction." *Working Paper*. February, 2005.
- Epple, Dennis and Glenn J. Platt. "Equilibrium and Local Redistribution in an Urban Economy when Households Differ in both Preferences and Incomes." *Journal of Urban Economics*. Vol. 43 (1998): 23-51.

- Epple, Dennis and Romer, Thomas. "Mobility and Redistribution." *Journal of Political Economy*. Vol. 99 (August, 1991): 828-858.
- Epple, Dennis and Holger Sieg. "Estimating Equilibrium Models of Local Jurisdiction." *Journal of Political Economy*. Vol. 107 (August, 1999): 645-681.
- Feenstra, Robert C., and James A. Levinsohn. "Estimating Markups and Market Conduct with Multidimensional Product Attributes." *Review of Economic Studies*. Vol. 62 (January, 1995): 19-52.
- Ferreira, Maria Marta. "Estimating the Effects of Private School Vouchers in Multi-District Economies." *Working Paper*. July, 2005.
- Geweke, John F. "Monte Carlo Simulation and Numerical Integration." In H. Amman, D. Kendrick and J. Rust (eds.), *Handbook of Computational Economics*. Amsterdam: North-Holland, 1996, 731-800.
- Griliches, Zvi. "Hedonic Price Indexes for Automobiles: An Econometric Analysis of Quality Change." *The Price Statistics of the Federal Government, 1961, General Series, No. 73* (New York: NBER). Reprinted in *Price Indexes and Quality Change*, Zvi Griliches ed. Harvard University Press. 1971.
- Graves, Philip E. and Donald M. Waldman. "Multimarket Amenity Compensation and the Behavior of the Elderly." *American Economic Review*. Vol. 81 (December 1991):1374-1381.
- Gorman, W.M. "A Possible Procedure for Analyzing Quality Differentials in the Egg Market." *Review of Economic Studies*. Vol. 47. (October, 1980): 843-856.
- Herzog Jr., Henry and Alan M. Schlottmann. "Labor Force Migration and Allocative Efficiency in the United States: the Roles of Information and Psychic Costs." *Economic Inquiry*. Vol. 19 (July, 1981): 459-475.
- Kanemoto, Yoshitsugu. "Hedonic Prices and the Benefits of Public Policies." *Econometrica*. Vol. 56. (July, 1988): 981-989.
- Klonsky, Karen, and Martin D. Smith. "Entry and Exit in California's Organic Farming Sector." *Economics of Pesticides, Sustainable Food Production, and Organic Food Markets*. edited by Darwin C. Hall and L. Joe Moffitt. Elsevier Science. 2002.
- Lancaster, Kelvin J. "A New Approach to Consumer Theory." *Journal of Political Economy*. Vol. 74. (April, 1966): 132-157.
- _____. *Variety, Equity, and Efficiency*. Columbia University Press. New York, 1979.

- Mäler, Karl-Göran. *Environmental Economics: A Theoretical Inquiry*. The John Hopkins University Press for Resources for the Future. 1974.
- Matzkin, Rosa L. "Nonparametric Estimation of Nonadditive Random Functions." *Econometrica*. Vol. 71 (September, 2003): 1339-1375.
- Murdock, Jennifer. "Handling Unobserved Site Characteristics in Random Utility Models of Recreation Demand." *Journal of Environmental Economics and Management*. Vol. 51 (January, 2006): 1-25.
- Nechyba, Thomas, J. "Existence of Equilibrium and Stratification in Local and Hierarchical Tiebout Economies with Property Taxes and Voting." *Economic Theory*. Vol. 10. (1997): 277-304.
- _____. "School Finance Induced Migration and Stratification Patterns: The Impact of Private School Vouchers." *Journal of Public Economic Theory*. Vol. 1. (1999): 5-50.
- _____. "Mobility, Targeting, and Private-School Vouchers." *American Economic Review*. Vol. 90. (March, 2000): 130-146.
- Ortalo-Magné, François and Sven Rady. "Heterogeneity within Communities: A Stochastic Model with Tenure Choice." *Working Paper*. February, 2006.
- O'Sullivan, Arthur, Terri A. Seton, and Steven M. Sheffrin. *Property Taxes and Tax Revolts: The Legacy of Proposition 13*. Cambridge University Press, 1995.
- Palmquist, Raymond B. "Estimating the Demand for the Characteristics of Housing." *Review of Economics and Statistics*. Vol. 66. (August, 1984): 394-404.
- Palmquist, Raymond B. and Adis Israngurka. "Valuing Air Quality with Hedonic and Discrete Choice Models." *American Journal of Agricultural Economics*. Vol. 81. (November, 1999): 1128-1133.
- Palmquist, Raymond B. "Weak Complementarity, Path Independence, and the Intuition of the Willig Condition." *Journal of Environmental Economics and Management*. Vol. 49 (January, 2005): 103-115.
- _____. "Property Value Models." in *Handbook of Environmental Economics, Volume 2*. edited by Karl Gören-Mäler and Jeffery Vincent. North Holland Press. 2005.
- Petrin, Amil. "Quantifying the Benefits of New Products: The Case of the Minivan." *Journal of Political Economy*. Vol. 110 (August, 2002): 705-729.

- Phaneuf, Daniel J. and V. Kerry Smith. "Recreation Demand Models," in *Handbook of Environmental Economics, Volume 2*. edited by Karl Göran-Mäler and Jeffery Vincent. North Holland Press. 2005.
- Polinsky, Mitchell A. "The Demand for Housing: A Study in Specification and Grouping." *Econometrica*. Vol. 45 (March, 1977): 447-462.
- Poterba, James M. "Housing and Taxation: Old Questions, New Answers." *American Economic Review*. Vol. 82. (May, 1992): 237-242.
- Provencher, Bill and Richard C. Bishop. "Does Accounting for Preference Heterogeneity Improve the Forecasting of a Random Utility Model? A Case Study." *Journal of Environmental Economics and Management*. Vol. 48 (January, 2004): 793-810.
- Rhode, Paul W. and Koleman S. Strumpf. "Assessing the Importance of Tiebout Sorting: Local Heterogeneity from 1850 to 1990." *American Economic Review*. Vol. 93 (December, 2003): 1648-1677.
- Ridker, Ronald. *Economic Costs of Air Pollution*. Frederick A. Praeger Publishers, 1967.
- Roback, Jennifer. "Wages, Rents, and the Quality of Life." *Journal of Political Economy*. Vol. 90 (December, 1982): 1257-1278.
- _____. "Wages, Rents, and Amenities: Differences Among Workers and Regions." *Economic Inquiry*. Vol. 26 (January, 1988): 23-41.
- Rosen, Sherwin. "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition." *Journal of Political Economy*. Vol. 82 (January, 1974): 34-55.
- _____. "Wage-Based Indexes of Urban Quality of Life." in P. Mieszkowski and M. Straszheim, eds., *Current Issues in Urban Economics*. John Hopkins University Press. Baltimore. 1979.
- Saez, Emmanuel. "Reported Incomes and Marginal Tax Rates, 1960-2000: Evidence and Policy Implications." *NBER Working Paper #10273*. January, 2004.
- Samuelson, Paul A. "Consumption Theory in Terms of Revealed Preferences." *Economica*. Vol. 15. (November, 1948): 243-253.
- Scotchmer, Suzanne. "Hedonic Prices and Cost/Benefit Analysis." *Journal of Economic Theory*. Vol. 37 (1985): 55-75.
- _____. "The Short-Run and Long-Run Benefits of Environmental Improvement." *Journal of Public Economics*. Vol. 30. (1986): 61-81.

- Sieg, Holger, V. Kerry Smith, H. Spencer Banzhaf and Randy Walsh. "Estimating the General Equilibrium Benefits of Large Changes in Spatially Delineated Public Goods." *International Economic Review*. Vol. 45. (November 2004): 1047-1077.
- _____. "Interjurisdictional Housing Prices in Location Equilibrium." *Journal of Urban Economics*. Vol. 52 (2002): 131-153.
- Smith, V. Kerry, and H. Spencer Banzhaf. "A Diagrammatic Exposition of Weak Complementarity and the Willing Condition." *American Journal of Agricultural Economics*. Vol. 86 (May, 2004): 455-466.
- Smith, V. Kerry, Holger Sieg, H. Spencer Banzhaf, and Randy Walsh. "General Equilibrium Benefits for Environmental Improvements: Projected Ozone Reductions under EPA's Prospective Analysis for the Los Angeles Air Basin." *Journal of Environmental Economics and Management*. Vol. 47. (2004): 559-584.
- So, Kim S., Peter F. Orazem, and Daniel M. Otto. "The Effects of Housing Prices, Wages, and Commuting Time on Joint Residential and Job Location Choices." *American Journal of Agricultural Economics*. Vol. 83. (November, 2001): 1036-1048.
- Starrett, David A. "Land Value Capitalization in Local Public Finance." *Journal of Political Economy*. Vol. 89 (April, 1981): 306-327.
- Taylor, Laura O. "The Hedonic Method," in *A Primer on Nonmarket Valuation*, edited by Patricia A. Champ, Kevin J. Boyle, and Thomas C. Brown. Kluwer Academic Publishers. 2003.
- Tiebout, Charles M. "A Pure Theory of Local Expenditures." *Journal of Political Economy*. Vol. 64 (October, 1956): 416-424.
- Timmins, Christopher, and Jennifer Murdock. "A Revealed Preference Approach to the Measurement of Congestion in Travel Cost Models." *Working Paper*. November, 2005.
- Topkis, Donald M. *Supermodularity and Complementarity*. Princeton University Press, Princeton, New Jersey, 1998.
- Varian, Hal R. "The Nonparametric Approach to Demand Analysis." *Econometrica*. Vol. 50. (July, 1982): 945-974.
- Walsh, Randall L. "Analyzing Open Space Policies in a Locational Equilibrium Model with Endogenous Landscape Amenities." *Working Paper*. 2003.

- Westoff, Frank. "Existence of Equilibria in Economies with a Local Public Good." *Journal of Economic Theory*. Vol. 14. (1977): 84-112.
- Wiatrowski, William J. "Changing Retirement Age: Ups and Downs." *Bureau of Labor Statistics Monthly Labor Review*. (April 2001): 3-12.
- Wu, JunJie and Seong-Hoon Cho. "Estimating Households' Preferences for Environmental Amenities Using Equilibrium Models of Local Jurisdictions." *Scottish Journal of Political Economy*. Vol. 50 (May 2003): 189-206.

II. Data Sources

Alexis, Andy and Paul Cox. *The California Almanac of Emissions and Air Quality: 2005 Edition*. California Air Resources Board, Planning and Technical Support Division. 2005.

Bolda, Mark, Laura J. Tourte, Karen M. Klonsky, and José E. Bervejillo. "Sample Costs to Produce Organic Strawberries: Central Coast Region." California Cooperative Extension, 2003.

Bolda, Mark, Laura J. Tourte, Karen M. Klonsky, and Richard L. De Moura. "Sample Costs to Produce Strawberries: Central Coast Region." California Cooperative Extension, 2004.

California Agricultural Statistics Service, *Summary of County Agricultural Commissioners' Reports*, 1999.

California Department of Finance, *California Statistical Abstract*, 2001.

California Air Resources Board, *Hourly Ozone Data*.

California Department of Education, *Academic Performance Index*, Annual Data.

California Employment Development Division, Labor Market Information Division, *Occupational Employment Statistics Survey*, Annual Data.

Dataquick, *Assessor Data from California Counties*.

Federal Home Loan Mortgage Corporation, *30-year Fixed Rate Mortgage Data*.

Metrolist Multiple Listings Service, *Online Data on Homes Currently for Sale*, October 2005. <http://www.metrolistmls.com/>.

Sacramento Bee Newspaper, Online Archive of Recent Housing Sales in the Sacramento Area, October 2005, <http://ssl.sacbee.com/onboard/homes.html>.

United States Department of Commerce, Bureau of the Census, *County-to-County Migration Patterns*, 2000.

United States Department of Commerce, Bureau of the Census, *Tiger Street Maps*, 2000.

United States Department of Commerce, Bureau of the Census, *County-to-County*

Commuting Patterns, 2000.

United States Department of Commerce, Bureau of the Census, *Census Transportation Planning Package Special Tabulation, 2000.*

United States Department of Commerce, Bureau of the Census, *Equal Employment Opportunity Special Tabulation, 2000.*

United States Department of Labor, Bureau of the Labor Statistics, *Consumer Price Index for Housing, 1995-2005.*

Appendices

Appendix A: Proof of Theorem 1

Given non-wage income, the first-stage parameter estimates, and values for the CES parameter and the composite unobserved public good, the estimator requires that there exists a set of values for the heterogeneous parameters, $A_z(\alpha, \gamma, \theta)$, that rationalize each observed location choice. This is stated formally as (A1).

$$(A1) \quad \text{For } z = 1, \dots, Z, \exists A_z(\alpha, \gamma, \theta) \text{ s.t. :}$$

- (i) $A_z(\alpha, \gamma, \theta) \subset \mathfrak{R}_+^{N+3}$
- (ii) $A_z(\alpha, \gamma, \theta) \neq \emptyset$
- (iii) For working households: $\theta_1 > 0$, $\theta_2 > 0$, and
for all households: $\alpha > 0$, $\sum \gamma = 1$
- (iv) $V_z[A_z(\alpha, \gamma, \theta) | \hat{y}, \rho, \delta, \xi] > V_q[A_z(\alpha, \gamma, \theta) | \hat{y}, \rho, \delta, \xi]$
 $\forall q = 1, \dots, Z, \text{ where } q \neq z.$

Given $\eta < 0$, $\nu > 0$, $\beta > 0$, there is guaranteed to be some threshold for the bandwidth on the kernel estimation of the composite unobserved public good, above which (A1) must hold. The proof of this statement proceeds in three stages. First, lemma 1 demonstrates there must be some allocation of the composite unobserved public good, $\tilde{\xi}$, which is strictly increasing in the price of housing and guarantees that (A1) holds for all retired households⁸⁸. Second, lemma 2 demonstrates that (A1) must hold for all working households at the $\tilde{\xi}$ allocation from lemma 1. Finally, theorem 1 explains that as the bandwidth on the kernel estimator of ξ increases, there must be some threshold above which the estimated allocation is monotonically increasing in prices. Because ξ is only defined up to a monotonic

⁸⁸ To avoid confusion with the notation, the bar superscripts on the composite unobserved public good are suppressed in this appendix.

transformation, the estimated allocation can be rescaled to the $\tilde{\xi}$ allocation that satisfies lemmas 1 and 2, completing the proof.

Lemma 1: For retired households with income and relative preferences defined by $(y \geq 0, \gamma_\xi = 1)$, there exists an allocation of the composite unobserved public good, $\tilde{\xi}$, such that (A1) is satisfied and $\tilde{\xi}$ is monotonically increasing in the price of housing.

Proof of Lemma 1:

For a retired household the choice set is restricted to housing communities. Order the J communities by their price of housing: $p_1 < p_2 < \dots < p_J$. Now let ${}_j\alpha_{j+1}$ denote a boundary indifference point for communities j and $j+1$, such that $V_j({}_j\alpha_{j+1}) = V_{j+1}({}_j\alpha_{j+1})$. Equation (A2) uses the indirect utility function in (3.1) to express ${}_j\alpha_{j+1}$ as a function of prices, income, the unobserved public good, and the homogeneous structural parameters, where $g_j = \xi_j$ because $\gamma_\xi = 1$.

$$(A2) \quad {}_j\alpha_{j+1} = \frac{X_j - X_{j+1}}{\xi_{j+1}^\rho - \xi_j^\rho}, \quad \text{where } X_j = \left[\exp\left(\frac{y^{1-\nu} - 1}{1-\nu}\right) \exp\left(-\frac{\beta p_j^{\eta+1} - 1}{\eta+1}\right) \right]^\rho.$$

Normalize the unobserved public good in community 1 such that $\xi_1 = 1$ and choose a larger but finite value for ξ_2 . Given $\xi_1 < \xi_2$, the assumed signs of β, η, ν imply $0 < {}_1\alpha_2$. Now, there must be some finite value for ξ_3 such that $\xi_2 < \xi_3$ and $0 < {}_1\alpha_2 < {}_2\alpha_3$ because in general ${}_j\alpha_{j+1} \rightarrow \infty$ as $\xi_{j+1} \rightarrow (\xi_j)^+$. This logic can be applied repeatedly to choose values for $\xi_4 < \dots < \xi_J$ such that ${}_1\alpha_2 < {}_2\alpha_3 < \dots < {}_{J-1}\alpha_J$. Within this monotonic ordering of boundary indifference points, community j is the utility maximizing choice for all

$\alpha : {}_{j-1}\alpha_j < \alpha < {}_j\alpha_{j+1}$. This follows from the observation that $V_j(\alpha) < V_{j+1}(\alpha)$ if $\alpha > {}_j\alpha_{j+1}$.

Thus, the constructed vector $\tilde{\xi} = 1 < \xi_2 < \dots < \xi_j$ guarantees that (A1) holds for retired households with $(y \geq 0, \gamma_\xi = 1)$. Furthermore, the ranking of communities by the unobserved public good matches the ranking by housing prices.

QED.

Lemma 2: Given the $\tilde{\xi}$ allocation from lemma 1, (A1) is satisfied for working households with non-wage income and relative preferences defined by $(\hat{y} \geq 0, \gamma_\xi = 1)$.

Proof of Lemma 2:

Consider a working household in any (j, k) location with non-wage income and relative preferences defined by $(\hat{y} \geq 0, \gamma_\xi = 1)$. Given the household's set of wage-commute options, let $\theta_2 = 0$ and $\tilde{\theta}_1 > 0$ be the solution to (A3).

$$(A3) \quad w_k = \tilde{\theta}_1 \max_{m \neq k} \{ w_m \}.$$

Any household with $\theta = (\tilde{\theta}_1, 0)$ will maximize its virtual wage income by commuting to job location k . This is equivalent to a situation where the household has exogenous income of $y = \hat{y} + w_k$. The $\tilde{\xi}$ allocation from lemma 1 defines a range of values for α that rationalize the choice of community j for such a household. Let α_j^* represent the midpoint of that range. Now let $\tilde{\theta}_2 > 0$ represent the shadow value of commute time that would decrease utility at location (j, k) to the level of utility at the second best location. That is, $\tilde{\theta}_2$ satisfies (A4).

$$(A4) \quad V_{j,k}(\hat{y}, \alpha_j^*, \gamma_\xi = 1, \tilde{\theta}_1, \tilde{\theta}_2) = \max_{l \neq j} \left\{ V_{l,k}(\hat{y}, \alpha_j^*, \gamma_\xi = 1, \tilde{\theta}_1, 0) \right\}.$$

It must be the case that $j, k = \max_{l,m} \left\{ V_{l,m}(\hat{y}, \alpha_j^*, \gamma_\xi = 1, \tilde{\theta}_1, \tilde{\theta}_2) \right\}$ because the indirect utility function is strictly increasing in income and $t_{j,k} > 0 \quad \forall j, k$.

QED.

Theorem 1: As the bandwidth on the kernel estimator of ξ increases there must be some threshold above which the predicted allocation of ξ guarantees that (A1) holds for any level of non-wage income, assuming all variables are finite with $\eta < 0$, $\nu > 0$, $\beta > 0$, and $(P, w, t, y) > 0$.

Proof of Theorem 1:

The distribution of ξ is estimated from the distribution of housing prices, conditional on observed public goods: $F(\xi) = F_{p|g=g_j}(p)$ in equation (18). As the bandwidth on the kernel estimator of this distribution function increases, $\hat{F}_{p|g=g_j}(p)$ converges to a line with a positive slope because $(\partial F_{p|g=g_j}(p) / \partial p) > 0$ by construction. Given finite values for prices and the observed public goods, there must be some finite threshold on the bandwidth above which $\hat{F}(\xi)$ is monotonically increasing in the price of housing. Because ξ is normalized to equal its distribution function the predicted allocation, $\hat{\xi} = \hat{F}(\xi)$, must be monotonically increasing in the prices. Finally, since $\hat{\xi}$ has no natural units, it can always be rescaled to equal the $\tilde{\xi}$ allocation from lemmas 1 and 2.

QED.

Appendix B: Genetic Algorithm

The *gloop_genetic* algorithm starts by choosing a random initial population of points from a uniform distribution defined by the starting solution space. Call this generation 0. Then the user-supplied function, f , is evaluated at every point in the initial population, and the population of points are ranked according to their function values. The initial set of points survive, reproduce, and mutate to create the next generation of points, generation 1. This process repeats to create each subsequent generation until one of the stopping criteria is met.

At the end of each generation a fixed number of the highest ranked points is chosen to survive without alteration to the next generation. The fixed number of survival points is chosen by the user.

Simulating reproduction and mutation for generation 1 requires choosing “parents” from the initial population of points. This is done by creating a cumulative distribution function based on the discrete set of points in the initial population and inverting the CDF to choose parents. The probability region attached to each point is proportional to the square root of its ranking as shown in equation (A5). Thus, higher ranking points have higher probabilities of being selected to produce the next generation. In the equation, P refers to the probability that an individual point is chosen, N to the size of the population, R to an individual point’s ranking within the population, and C to the unique constant that guarantees both restrictions hold.

$$(A5) \quad P_i = C\sqrt{R_i}, \text{ such that } \sum_i^N C(\sqrt{R_i}) = N \text{ holds}^{89}.$$

A fixed number of parents are chosen from the inverse CDF, with replacement. Individuals are mutated by adding a random disturbance chosen from a multivariate normal distribution centered on 0. Reproduction occurs in two ways; through *DNA exchange* and *hypercube mating*. The degree of mutation relative to reproduction, the standard deviation of

⁸⁹ This particular concept of “fitness scaling” is common to many genetic algorithms, including the one contained in the *Mathworks* global optimization toolbox, as described in the *Mathworks Genetic Algorithm and Direct Search User’s Guide, Volume One*.

the multivariate normal distribution, and the balance between DNA and hypercube mating are determined by the user.

DNA exchange involves two points exchanging coordinates. For each pair of points, DNA exchange requires a random vector of zeros and ones. The elements of the random vector are used as dummy variables to select individual genes from both parent points to create a child point. For example, define two points x and y , and a random vector as follows:

$$(A6) \quad x = [x_1 \ x_2 \ x_3 \ x_4], \quad y = [y_1 \ y_2 \ y_3 \ y_4], \quad \text{rand}=[0 \ 1 \ 0 \ 1].$$

The new point created through DNA exchange is $z=[y_1 \ x_2 \ y_3 \ x_4]$.

Hypercube mating takes two parents and uses a random draw to select a new point somewhere in the hypercube defined by the parents. It is important to note that reproduction will never produce points outside the initial solution space defined by the user. This can occur through mutation however.

The degree of mutation is determined by the user. For each parameter that the algorithm is searching over, *gloop_genetic* calculates the standard deviation from the current generation of random points. Then that standard deviation is multiplied by an inflation factor specified by the user. The result is used to take random draws (or mutations) from an independent multivariate normal distribution centered on zero. The draws are added to existing parent points to create mutated children. If the user requests persistent mutation, the inflation factor*st.dev.(gen 0) is applied to every generation, and the degree of mutation does not diminish over time. Conversely, if the user requests degenerative mutation the inflation factor*st.dev.(gen i) is applied to every generation so that the degree of mutation decreases as the population starts to converge on a solution.

There are five key advantages of *gloop_genetic*. The first is that it is a non-derivative method. The user does not need to provide a derivative of the function defined by f . The second advantage is that the algorithm is adaptive. In general, it samples areas of the solution space with higher function values more thoroughly. To follow the natural selection analogy, higher points are more likely to survive and reproduce. The third advantage is that

the algorithm may find an optimal solution outside the initial solution space. The fourth advantage is that it works with vectorized functions. That is, if the user-specified function can accept a matrix of points and return a vector of solutions, *gloop_genetic* can work faster by calculating the values for an entire generation of points in a single step without requiring loops. Finally, the fifth advantage is that the algorithm does not require much memory because it only stores the current generation in memory; previous generations are not stored in the computer's memory.