5-2010

Three Essays on Environmental- and Spatial-Based Valuation of Urban Land and Housing

Lu Liu
Utah State University

THREE ESSAYS ON ENVIRONMENTAL- AND SPATIAL-BASED
VALUATION OF URBAN LAND AND HOUSING

by

Lu Liu

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

of

DOCTOR OF PHILOSOPHY

in

Economics

Approved:

Paul M. Jakus Major Professor	Arthur J. Caplan Committee Member
Reza Oladi Committee Member	Frank Caliendo Committee Member
Mark Brunson Committee Member	Byron R. Burnham Dean of Graduate Studies

UTAH STATE UNIVERSITY
Logan, Utah

2010

ABSTRACT

Three Essays on Environmental- and Spatial-Based Valuation

of Urban Land and Housing

by

Lu Liu, Doctor of Philosophy

Utah State University, 2010

Major Professor: Dr. Paul Jakus
Department: Applied Economics

This dissertation attempts to provide a comprehensive examination on the non-market valuation of the effect of open space amenities and local public infrastructure on the value of urban land and housing with both spatial heterogeneity and project heterogeneity. The demand for raw land is a derived demand for housing built on it. Therefore, we need to examine the land market and the housing market together. On the one hand, we estimate the value of urban land in a market that does not satisfy the usual assumptions of a competitive market structure as well as incentive incompatibility issues for transaction participants, with an application to a Chinese regional wholesale land market. These two violations to the traditional hedonic theory also generate two separate valuations on land with differentiated characteristics. On the other hand, we utilize the relative plane coordinates system, the three-dimensional distances, as well as the aggregate weight matrix, to implement the spatial hedonic estimation on the high-rise residential buildings in the same regional housing retail market in China. After these two steps, this dissertation, therefore, focuses on the profit maximization behavior of the property developer, which is the key role to link the factor market (i.e., the land market) and the commodity market (i.e., the housing market) together. Two methods are then employed to implement the hypothesis test on the hedonic price estimation including both inputs and outputs. First, a set of partial derivatives of the profit function with respect to various characteristics gives us the relationship between the marginal valuations in the land market and in the housing market. Second, we introduce a joint estimation approach that we call the spatial full information maximum likelihood (SFIML), which considers the land market, the housing market, and the property developer's profit maximization behavior all together in the estimation. Finally, we conduct a hypothesis test in both of these two scenarios to examine the validity of our linked markets assumption on the hedonic price estimation.

(175 pages)

To my parents, who brought me life;

To my wife, who brightened it.

ACKNOWLEDGMENTS

I am deeply grateful to my parents and my wife; without their emotional support during my four years of hard study in the US, this dissertation would not have been possible. I am also deeply grateful to the faculty of the Department of Applied Economics (and former Department of Economics) at Utah State University, who brought me through the threshold of modern economics and gave me rigorous training that was necessary for me to conduct independent economic research in my future career; also without their financial support during the four years, this dissertation would not have been possible. Dr. Paul Jakus's thoughtfulness has been inspiring. As my major professor, his patient guidance and time commitment during the last three years (the three fourth of my entire study in the program) have trained me with necessary econometric skills to write this dissertation and definitely enhanced its quality. The cooperation with him in the OHV project has not only given me experience in professional economic research, but has also brought me the first publication in a refereed English journal in my life. His seemingly unlimited enthusiasm in research inspired and encouraged me on the road of economics too. In addition, without his detailed suggestions and editing, this dissertation would not have been possible as well. Dr. Arthur Caplan, as my advisor in the first year of the PhD program, also as the teacher of my Micro 1 and Environmental Economics, has trained me with the solid theoretical foundation, and also inspired me on the rigorous attitude in research. I would never forget that, in the summer of the first year, he sent me comments on my draft paper even during his stay in Africa. In addition, his detailed editing comments on my dissertation are appreciated as well. Dr. Reza Oladi, as the teacher of my Micro 3 and Trade, has not only strengthened my theoretical foundation, but also pushed me into the field of game theory and open economy. The cooperation with him on research was a nice experience in my program as well. His guidance on how to publish a paper will benefit me for a long time. Dr. Frank Caliendo, as my Macro 3 teacher, has helped me in the dynamic field. His seminar-like classes have inspired me on the research of life-cycle theory to a great extent. Although such topics did not become my dissertation topic, the methods and skills that I learned will sure be beneficial to my career. Dr. Mark Brunson, as the outside member of my committee, has given me detailed suggestions on my proposal and dissertation. His insight and expertise have inspired me on my dissertation from outside but related to economics. Besides these five dissertation committee members, I would also like to thank two professors, Dr. Kenneth Lyon and Dr. Basudeb Biswas, who taught me four courses each. Without their rigorous training, I would not have the basic knowledge and skills to conduct further research. In a word, my four years of life and study at USU are a treasure to my entire life, which not only gives me the skills to "go fishing," but also mentally inspires and shapes me. The professors are my teachers, but also my friends.

In addition, I would also like to thank my former company and colleagues in my hometown, Chengdu SAGA Organization Ltd., without whose generous support on the housing retail sales data, part of my dissertation would not have been possible. Also, the two-year work experience with them has given me the first-hand perception about the market, which inspired me on a series research of land and housing.

Finally, I would like to thank my former university in China, Southwestern University of Finance and Economics, as well. Without seven years of study there, I would not even know what economics is, not to mention my great interest in it.

Lu Liu

CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

It is commonly acknowledged that modern hedonic theory should be credited to Rosen (1974), who proposed an equilibrium model of product differentiation. The hedonic approach has seen widespread applications to help value: air quality, open / green space, public transportation, water proximity and quality, and planned local infrastructure. Traditional hedonic theory relies upon two critical assumptions: the competitive market structure and "matching" property prices with the market participants' true valuations. If the market, however, is characterized by a monopolistic seller, then we do not have a set of offer curves as the traditional hedonic theory predicts. Instead, we end up with only one offer function which stands alone in the market. In addition, in some special cases such as an English auction setting, the actual sales price may not represent both the seller's and the buyer's true valuations since a possible auction premium may exist. In either case, we cannot directly use the observed sales price to estimate the hedonic equilibrium and implicit marginal prices. Under these market conditions, the sales price fails to represent the true valuation of the market participants, due to the monopolistic seller and the incentive incompatibility issue in the English auction. To our knowledge, no study has been done to examine the hedonic valuation when confronting these two violations to traditional assumptions of hedonic theory.

In Chapter 2, data on the land market in China provides us with an opportunity to examine the two violations. The Chinese regional land market is characterized by a monopolistic land seller (the local government) and multiple buyers (developers) who purchase land via English auction. We are able to take advantage of these market features in two ways. First, the "asking price" of the government seller is used to derive its true valuation, so that one can estimate the offer function of the monopoly seller. On the buyer's side, the winning bid does not necessarily reflect the true valuation of the buyers. But with the known asking price and winning bid, the incentive incompatibility properties of the English auction can be exploited to recover the true valuation of the buyers.

Our empirical analysis looks at the marginal implicit values for characteristics of raw, developable land. The characteristics considered include development restrictions regarding housing density and minimum green space, *in situ* and planned infrastructure such as parks and public transportation systems, and neighborhood effects. Because no equilibrium price function exists, we conduct the analysis separately for the land seller (the local government) and buyers (land developers). We find that, contrary to standard hedonic theory, the marginal implicit characteristics are not equal across buyers and sellers.

The natural extension of the study in Chapter 2 is to examine the retail housing market, i.e., look at the hedonic equilibrium in the structures built upon the raw land considered in Chapter 2. The Chinese regional housing market consists of housing units in different housing projects. Unlike the relatively "sparse" residential development pattern common in the US and other countries, the style of residential development in China is more concentrated and dense. In fact, many large cities in Asia develop in a similar manner, and their residential buildings have the "high-rise" shape. Over the past 20 years, high-rise residential development has expanded from the coastal region to the inland region, and it is currently the prevalent urban-development pattern in China. The high-rise residential pattern challenges the traditional spatial hedonic techniques because the standard two-dimensional concept in space does not fit the situation well. To our knowledge, no study has been done to conduct the hedonic estimation with respect to the high-rise residential pattern.

We adapt our spatial econometric model to reflect the potential for three-dimensional spatial relationships within a high-rise apartment complex, as well as the two-dimensional spatial relationships between complexes. Our equilibrium hedonic price function explains apartment sales prices as a function of project-specific attributes such as housing density and *in situ* and planned infrastructure such as parks and public transportation, and apartment-specific characteristics such as the size of the apartment and the floor on which it is located.

While Rosen (1974) and many subsequent studies have focused on the different characteristics of the output, Palmquist (1989) extends the study into the differentiated factors of production with a focus on land. Palmquist treats land as a differentiated production input, and assumes that, this differentiated factor (land, in this example) is purchased by a buyer following a derived demand for the input. To our knowledge, while most previous hedonic studies focus on the "final product" (retail housing), the critical role of the property developer has long been ignored. In fact, it is the property developer that links the land and housing markets together. Although the studies of Palmquist (1989) and Wu (2006) (among others) have shed light on the theoretical link between the factor market and commodity market, to our knowledge no study has attempted to empirically link the derived demand for land to the supply of retail housing.

Chapter 4, therefore, focuses on the profit maximization behavior of the property developer. The property developer is assumed to earn a positive profit from the English auction where the raw land parcel is traded with the local government, besides the common competitive market assumption. This profit arises from the premium due to the incentive incompatibility problem with the English auction, since the winner only needs to pay the amount at which the second highest bidder quits. With the developer's true valuation of land derived from Chapter 2, we test whether the parameters from the derived demand are consistent with those of the supply. Both separate estimation and joint estimation approaches are employed in the empirical models. A set of partial derivatives of the profit function with respect to various characteristics gives us the relationship between the marginal valuations in the land and housing markets, which then present a link between the estimation parameters in these two markets, and could be considered as constraints in the estimation parameters.

We also use a joint estimation approach that we call the spatial full information maximum likelihood (SFIML), which considers the land market, the housing market and the property developer's profit maximization behavior all together in the estimation. We use the results in the corresponding separate estimation in the housing market as the constraint on the SFIML parameters.

The results of the separate estimation model reject the null hypothesis that the calculated constraints are valid. In contrast, the joint estimation model fails to reject the null hypothesis, which provides a positive signal confirming the theoretical linkage in the hedonic price estimation.

References

Palmquist, R. B., 1989. Land as a differentiated factor of production: a hedonic model and its implications for welfare measurement. Land Economics 65 (February), 23-28.

Rosen, S., 1974. Hedonic prices and implicit markets: product differentiation in pure competition. Journal of Political Economy 82 (1), 34–55.

Wu, J., 2006. Environmental amenities, urban sprawl, and community characteristics. Journal of Environmental Economics and Management 52 (2) (September), 527-547, doi: 10.1016/j.jeem.2006.03.003.

CHAPTER 2

A SPATIAL HEDONIC STUDY FOR MONOPOLY SUPPLIED URBAN LAND VIA ENGLISH AUCTION: A CASE STUDY OF CHENGDU, CHINA

Abstract

This study estimates the effect of open space and local public infrastructure on the value of urban land in a market that does not satisfy the usual assumptions of the traditional hedonic theory. Our study uses data obtained from a Chinese regional land market characterized by a monopolistic land seller (the local government) and multiple buyers (developers) who purchase land via English auction. The "asking price" of the government seller is used to derive its true valuation, so that one can estimate the offer function of the monopoly seller. For developers, the winning bid does not necessarily reflect the true valuation of the buyers due to the incentive incompatibility properties of the English auction. Following Paarsch (1997), we use the difference between the asking price and the winning bid to calculate a "bid premium." The premium is then used to recover the distribution of the buyer's true valuation. Our estimates thus reveal the true marginal valuation for amenities and infrastructure associated with a given property by both buyers and sellers which, under our market conditions need not be equal because the usual hedonic equilibrium does not apply. In our study of land sold for residential development in Chengdu, China, we find that the seller and buyers differ in the marginal valuation of these land

characteristics. In addition, our study can be used to shed light on the "land financing" issue in China, land sales are a primary tool of local public financing.

1. Introduction

This study estimates the effect of open space and local public infrastructure on the value of urban land in a market that does not satisfy the usual assumptions of the traditional hedonic theory. Our study uses data obtained from a Chinese regional land market characterized by a monopolistic land seller (the local government) and multiple buyers (developers) who purchase land via English auction. We take advantage of these market features in two ways. First, the "asking price" of the government seller is used to derive its true valuation, so that one can estimate the offer function of the monopoly seller. On the buyer's side, the winning bid does not necessarily reflect the true valuation of the buyers. Given the known asking price and winning bid, the incentive incompatibility properties of the English auction can be exploited to recover the true valuation of the buyers. The implicit prices of the offer and bid functions reveal the true marginal valuation for amenities and infrastructure associated with a given property. In our study of land sold for residential development in Chengdu, China, we find that the monopolistic seller and buyers differ in the marginal valuation of these land characteristics.

The paper proceeds as follows: first, we briefly review the traditional hedonic theory under standard theoretical assumptions, along with a review of how this model has been applied to the valuation of open space amenities and infrastructure. After discussing the land market in China and the City of Chengdu, we present our data. We then discuss the properties of an English auction and its application to our study. Our empirical section consists of three parts: a Tobit model to estimate the auction premium paid by the buyers and its subsequent transformation into the distribution of the buyers' valuation, an empirical model for the monopolist's offer function, and, finally, a model of the bid function using the land buyers' derived true valuation.

2. Literature Review

It is commonly acknowledged that hedonic theory should be credited to Rosen (1974), who proposed an equilibrium model of product differentiation. In a competitive market setting, goods are assumed to be valued for their utility-bearing characteristics, and the interactions of buyers and sellers over multiple attributes yield the hedonic equilibrium price function. The hedonic price function is an envelope of the tangent points between the offer functions and the bid functions. The hedonic approach has been seen widespread application so, for this study, we will initially focus our literature review on open / green space, public transportation, water proximity, and planned local infrastructure. There are many studies that have shed light on our research, some of which we review below.

Anderson and West (2006) estimate the effects of proximity to open space on home sales price in the Minneapolis–St. Paul metropolitan area. They measure the size of the nearest amenity of different types in acres, such as neighborhood park, special park, golf course, cemetery, lake. Although many recent studies measure the total quantity of open space surrounding a home within a given distance or at multiple scales, they prefer to use the distance to the nearest open space, since they include the block group fixed effects, and homes in the same census block group often have the same overall pattern of surrounding land use. They also calculate the distance from each home to the nearest CBD. Again, home value is regressed on structural attributes, neighborhood characteristics as well as location, and environmental amenities. Census block group data are used as control variables. A log-log functional form is used in the estimation, with results showing that the value of proximity to open space is higher in neighborhoods that are characterized as: dense, near the CBD, high-income, high-crime, or home to many children. Anderson and West also find that the sales price of an average home increases with the proximity to neighborhood parks, special parks, and golf courses. However, they find that these results are sensitive to the inclusion of local fixed effects.

Asabere and Huffman (2009) measure the relative impacts of trails, greenbelts, and the interaction of trails with greenbelts on home values for over 10,000 sales of residential property occurring in and around Bexar County, Texas. A distinct feature of their study is that they use dummy variables to denote almost all the open space variables such as presence of a trail in the neighborhood, a greenbelt in the neighborhood, both trail and greenbelt, a golf course, a playground, tennis court, and a swimming pool. Actual distances from trail or greenbelt are measured based on the MLS database. In addition, they also consider additional sales-related variables including time-of-sale in sequential months, and type of financing (conventional versus others). A semi-log functional form is used in the estimation. Their study shows that trails, greenbelts, and trails with greenbelts are associated with roughly 2%, 4%, and 5% price premiums, respectively. The authors, therefore, confirm that the home value would be further enhanced when greenbelts are used to buffer trails and hence create greenways.

Bolitzer and Netusil (2000) examine the net effect of open space proximity on a home's sale price in urbanized Portland. They include all publicly owned open spaces and those privately owned large open spaces that exceed 10 acres. Public parks make up the majority of open spaces in this study. Proximity to an open space, open-space type and distance from the house to the central business district are obtained using a geographic information system (GIS) database. An "open space" dummy variable was created to reflect the presence of any open space within 1500 feet of a home. The sales price of a home is then regressed on structural characteristics, environmental characteristics, open space characteristics, and other neighborhood characteristics. Both linear and semi-log functional forms are used in regression and the results from the semi-log specification are preferred. Their results show that proximity to an open-space of certain type can have a positive and significant effect on a home's sale price in their study area, but they do not find that the negative externalities associated with open space adjacency dominate the positive externalities (as was found in other empirical studies).

Geoghegan et al. (1997) include two ecological landscape indices (diversity and fragmentation) to hedonic valuation on land use. In their study, they introduce a diversity index based on Shannon index and a fragmentation index which is the perimeter to air area ratio, fractal dimension (edge to interior) and the edge length between land use. They measure the two ecological indices at both a 0.1km and 1.0 km radius surrounding each housing transaction to capture the scale issue. Besides this buffer, they also consider structural characteristics (age of house, type of construction material, lot size, and whether lot is waterfront or not), locational characteristics (i.e., distance to the central business district, CBD), and accessibility (the distance to the nearest major road). Their study area is the 30-mile radius of the Washington DC, which they think is the maximum possible commute range of the market. Both census data on ethnic composition and income and GIS data on streets, highways, and hydrological systems are used. Without doing a tedious process of address-matching, they use a 3000 ft by 4000 ft size grid, and then geo-code them into GIS. In addition, they use dummy variable to capture differential tax rates and public services. In their regression, natural log functional form is used. Their research has found that for a smaller buffer, the marginal contribution of more open space is both positive and significant; while for a larger buffer, the effect is both negative and significant.

Irwin (2002) addresses the identification problems in a hedonic pricing model due to the endogenous explanatory variables, spatial error autocorrelation and multicollinearity. She distinguishes six types of open space by individuals' perceptions of neighboring open space, namely whether it is in a preserved state or developable. She also divides the open space into land that could be developed at anytime (cropland, pasture, or forest) versus land that has been permanently preserved in some way (privately owned land whose development rights have been sold or land that is publicly held). Irwin considers land ownership and land use as well, using a 400-meter radius around residential parcels as the study area. She also considers the proportion of neighboring land that is in low, medium, and high density residential development and commercial or industrial land use to capture the externality effects of neighboring development. Distance to the two major centers in the study area, i.e., Washington, DC and Baltimore is measured along major roads. A dummy variable is used to denote whether a residential property is located within one mile of the airport to examine the noise disamenity as well. In addition, several socioeconomic variables from the 1990 U.S. Census of Population measuring at the block group level and dummies for three of the four counties in the study area are also included.

In Irwin's study, residential sales price is regressed on structural characteristics associated with the house, neighborhood / locational variables, as well as neighborhood land use variables. Irwin compares log-log, semi-log functional forms, and a linear version of the Box-Cox transformation. Results show that the log-log and semi-log specifications do a better job than the linear model and a slight preference is given to the log-log model by ordinary least square estimation. Privately owned conservation lands, publicly owned conservation lands, nonmilitary open space have positive and significant effects on the value of neighboring residential properties relative to developable pasture land. Notably, Irwin randomly draws a subset of the data to control the inefficiency of the estimates caused by the remaining spatial error correlation. She first defines the nearest neighbors as parcels that are within 100 meters of each other and then uses 200, 400, and 600 meters of each other to test model robustness. She finds that the spillover effects from preserved open space are significantly greater than those associated with developable farmland and forest, and that pasture land generates a significantly greater spillover effect on residential property values than that of neighboring forests.

Leggett and Bockstael (2000) estimate the effects of water quality on residential land values along the Chesapeake Bay, in Anne Arundel County, Maryland. They use fecal coliform bacteria, which has serious human health implications, as a measure of water quality. They collect data for sales of waterfront property between July 1993 and August 1997 from the State of Maryland's Tax Assessment data base. Distance is measured from a parcel to the closest water quality monitoring stations. The authors calculate an inverse distance-weighted average of fecal coliform counts based on data from the nearest three monitoring stations for each waterfront property.

In addition, the appraised value of the structure by the tax assessors is also included in the regression. They include lot size and its square as explanatory variables as well. Commuting distances to the nearby cities (Annapolis, Baltimore, and Washington, DC) are measured using ARC/INFO software along road networks digitized in the Census Bureau's Tiger Line Files. Additional variables include black population as a percent of total population and percent of owner occupied housing in the Census block group.

In the regression, log-log, semi-log, inverse semi-log, and linear functional forms are compared. Leggett and Bockstael estimate two alternative dependent variables for each of the four specifications: one is market transaction price minus assessed value of the structure and, the other one is the market transaction price itself. The first one is explained as the "residual" land price. They use ordinary least squares to estimate all of the eight specifications, and find that both heteroscedasticity and spatial autocorrelation are in the OLS results. They argue that it is difficult to resolve these two problems at the same time, so they first focus on four specifications which do not exhibit heteroscedasticity and then re-estimate these specifications using spatial error model to correct spatial correlation. In the end, the inverse semi-log functional form is chosen to conduct a comparative study in welfare change. The model indicates that improvements in water quality can have a positive and significant effect on property values.

Lutzenhiser and Netusil (2001) estimate the effect of proximity to different open space types on a home's sale price in the city of Portland, Oregon. Open spaces are assigned to one of five categories: urban parks, natural area parks, specialty parks/facilities, golf courses, and cemeteries. Dummy variables were created to reflect the interaction between seven different zones that range in size from 200 to 300 feet and the open space types. Home prices are regressed on structural characteristics, environmental characteristics, neighborhood characteristics. The estimated effects are composed of three factors: the open space variable interacted with distance, and acreage and acreage squared interacted with open space type. Box-Cox transformation of the dependent variable is used in the estimation, where a maximum likelihood value for the parameter λ in the transformation is estimated. Their findings show that homes located within 1,500 feet of a natural area park, where more than 50% of the park is preserved in native and/or natural vegetation, have the largest increase in sale price. In addition, Lutzenhiser and Netusil show that natural area parks require the largest acreage to maximize sale price, and specialty parks are found to have the largest potential effect on a home's sale price.

Mahan et al. (2000) use the hedonic property price model to estimate the value of wetland amenities in the Portland, Oregon, for the metropolitan area with over 14,000 home sales records. Arc/Info GIS is used to generate the data, and wetland characteristics are based on the U.S. Fish and Wildlife Service's National Wetlands Inventory in Oregon. Their major land-cover categories include forested, scrub-shrub, emergent-vegetation, open-water wetlands, lakes and rivers or streams.

They record the size in acres of nearest wetland of any type (excluding lakes, rivers, and streams) and use a dummy variable to denote the type of nearest wetland. A raster system is used to calculate the Euclidean distance in feet from the centroid of the tax lot to the nearest edge of a feature, where all data are arranged in grid cells (52-feet square for each). They also measure the natural log of distance to the nearest open water linear wetland, water areal wetland, stream, river, lake, and improved public park. Housing prices are regressed on environmental amenities associated with a specific location, structural characteristics, neighborhood characteristics, and market segment variables. Notable neighborhood characteristics include the tax rate, distance to a central business district, a dummy variable for light traffic, elevation of property above sea level, slope of property as a percent, natural log of the distance in feet to nearest industrial zone, nearest commercial zone, and quality of view as indicated by county assessor (range 0-9, 0 if no view). Prices are logged in order to implement least squares regression in estimating the hedonic price function.

Two models are estimated based on different assumptions. In model 1, characteristics of the nearest wetland (size, distance, type) are assumed to affect property value; while in model 2, the distance to the nearest wetland of each type is assumed to influence property values. Their results show that increasing the size of the nearest wetland by one acre would increase a property's value by $24.39, while decreasing the distance to the nearest wetland by 1,000 feet would increase a property's value by $436.17. In addition, the type of wetland does not appear to matter to nearby residents.

Besides the literature that we have discussed above, some other notable examples of such studies include (but not limited to): Bates and Santerre (2001), Geoghegan (2002), Provencher et al. (2008), Sander and Polasky (2009), Schulz and Waltert (2009), and Shultz and King (2001). While most hedonic studies choose housing price as the research basis (i.e., the dependent variable in the regression), there are some studies that choose the value of land as the target variable. Since the structure of housing itself is an important factor that affects the housing price, for our study perhaps the value of raw land is a better basis for evaluating the open-space impact on property value. A good example is Cheshire and Sheppard (1995).

Cheshire and Sheppard (1995) estimate the capitalization of the value of the location-specific characteristics into land prices. Unlike the conventional approach which treats urban rent as the price of pure land, they argue that land itself is a composite good which embodies neighborhood, environmental characteristics and local public goods. They use data from Reading and Darlington during a comparatively stable period in the British housing market. The 1981 Census of Population is used to provide data of neighborhood characteristics. They also measure the accessibility of each house to the bus network as well as roads of different classes. They suggest that larger roads may increase the housing value since they provide better accessibility and more importantly, the possible conversion to commercial use. Accessible land amenity, non-accessible land amenity, percent of land in accessible open space, and percent of land in inaccessible open space are recorded in a 1 kilometer square around each structure.

Cheshire and Sheppard (1995) construct a very flexible land rent function, which uses an exponential form to regress the land rent on distance from town centre and angle of deflection from East. They suggest this form because they think it could allow for multiple radial asymmetries in land rents to emerge via the estimated parameters. This land rent function is then incorporated into the hedonic model where the Box-Cox functional form is used. The rental price is regressed on structural or location-specific characteristics, the quantity of land included with structure, set of indices of characteristics that are dichotomous, set of indices of characteristics that are continuously variable and the land rent function. One distinct feature is that they include the effect of closely correlated variables within one variable to resolve the colinearity between characteristics. They include both the congestible amenities and structure characteristics since they suggest that, in general they will not be correlated due to the "neighborhood" nature. Their findings show that, the rent does not monotonically decline from the CBD, but it increases in certain directions.

A few authors have found that proximity to public transportation or roads and highways can have a significant impact on property values. The effect is complex: good access to such infrastructure can make daily life more convenient, but it may also be associated with disameneties such as traffic noise and increased crime.

Gibbons and Machin (2003) evaluate the economic benefits of transport access, noting both the positive and negative impacts of proximity to a railway line. They distinguish between proximity to a railway line and the distance to a station to separate out environmental and transport access effects. Their research confirms that benefits of station proximity and high service frequencies are both capitalized in property prices. Nelson (1982) also reviews nine studies of the effect of highway noise, finding that highway noise levels decline to background levels within roughly 1,000 feet of a highway so that the effect on property values is contained to a relatively small segment of a market.

We now summarize the literature reviewed thus far. In these studies, open space has been interpreted very broadly as parks, wetlands, trails, rivers, creeks, or even unused land, and are normally measured in three ways. The first method uses only proximity, which is commonly calculated by the Euclidean distance (in feet or meters) from the centroid of the property to the nearest edge (or centroid as well) of a feature. The second method is to use dummy variables to show the existence of a feature within certain range of the property, e.g., within 100 feet, 1000 meters, and so on. The third approach is to combine a measure of proximity with a measure of size, where size of each feature is calculated by acreages or square meters. Other locational characteristics, such as distance to the central business district or employment centers in other nearby cities, are also frequently included. With regard to the functional form, it appears that the choice of functional forms is simply an empirical issue. Normally, linear, semi-log, and log-log functional forms are used and compared. Sometimes, Box-Cox transformation is also used to derive a more flexible functional form. Most studies use a combination of property sales data, GIS data (on streets and highways, hydrological systems, etc.), and the Census data (on both ethnic composition and income, etc.), which demonstrates the data requirements of hedonic studies. In regard to the valuation, sales price of a residential property is commonly regressed on structural characteristics, environmental characteristics, open space characteristics, and other neighborhood characteristic, as well as market segment variables.

While much of the hedonic literature uses a static approach, some hedonic studies involve data gathered over time. As Freeman (1993) has proposed, most environmental goods are time-variant and therefore may be lead to different price estimates over time. Riddel (2001) argues that if the time needed for full realization of amenity value is sufficiently long, then one should incorporate a time trend in the estimation. Common approaches to the time issue are to deflate sales price by some kind of housing price index (for example, Bolitzer and Netusil, 2000; Lutzenhiser and Netusil, 2001) or the consumer price index (for example, Geoghegan, 2002; Leggett and Bockstael, 2000). The choice of methods is, once again, an empirical issue.

As deflating by HPI appears to be one of the standard approaches, Diewert et al. (2010) argue that the housing price index needs to be decomposed into land and structure components, casting some light on the empirical difficulties of the prevalent use of HPI. In addition, the time-dummy method is also very popular in hedonic studies (see the discussion by Melser, 2005). For example, Provencher et al. (2008) include annual dummy variables to represent the temporal shifts in the residential property market.

In contrast to the previous studies, which focused on property values for already developed land, an important extension of Rosen's framework was presented by Palmquist (1989). Palmquist treats land as a differentiated production input and assumes that, this differentiated factor (land, in this example) is purchased by a buyer following a derived demand for the input. The supply side is similar to the Rosen's (1974) model, but Palmquist separates the characteristics vector into two parts: in addition to the usual assumption of exogenously determined characteristics, some characteristics could be endogenously determined by the buyer. The bid function for raw land hence arises from the derived demand for existing exogenous characteristics, as well as those characteristics that can be manipulated.

Traditional hedonic theory is based on two critical assumptions: the competitive market structure and the matching property prices with the market participants' true valuations. However, if the market is characterized by a monopolistic seller, then we do not have a set of offer curves as the traditional hedonic theory predicts. Instead, we end up with only one offer function which stands alone in the market (see Fig. 2.1). In addition, in some special cases such as an English auction setting, the actual sales price will not represent the market participants' true valuations since possible auction premium may exist (see Fig. 2.2). In these scenarios, we cannot directly use the observed sales price to estimate the hedonic price function because it fails to represent the true valuation of the market participants, due to the monopolistic seller and the incentive incompatibility issue for all the participants in the English auction. To our knowledge, no study has been done to examine sales of property when confronting these two violations to the traditional assumptions of the hedonic pricing theory.

3. Market Setting and Data

3.1. The Land Market in China

The land market in China provides us with an opportunity to examine the two violations to traditional hedonic theory mentioned above. In China, all land is owned by either the central government or local government, although the precise entity holding ownership is usually not specified. The sale of land for development is in essence a long term lease, with the term varying from 40 to 70 years. The maturity for residential use land is 70 years, which is a time period long enough to have generated an active real estate market for developers and private citizens seeking housing. Currently the most popular transaction method for private development in the "wholesale" land market is an auction. Two types of auction are used in the market: a Type 1 auction is held in an auction hall at a particular time, with the land sale completed later that same day. In contrast, a Type 2 auction

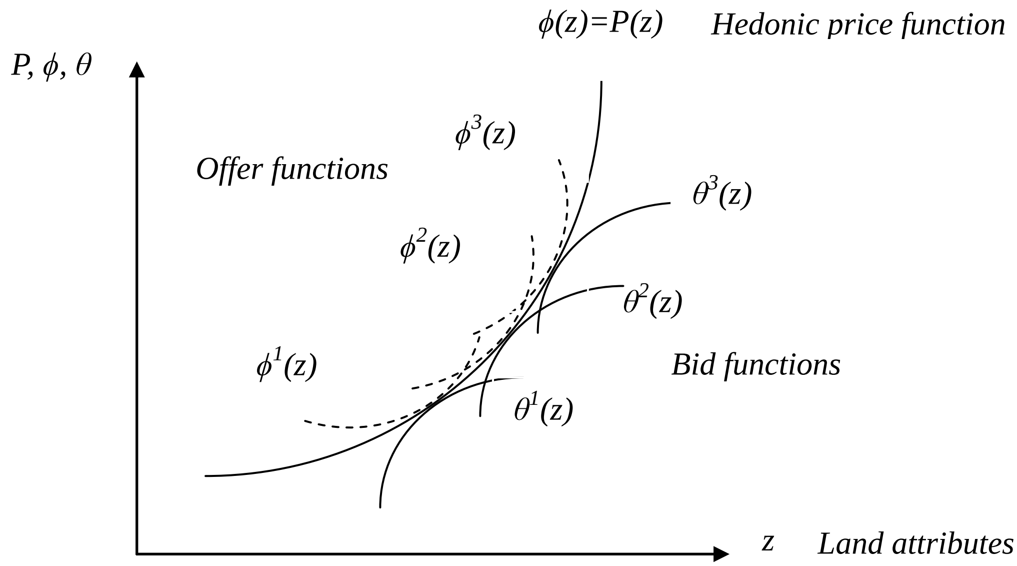

Fig. 2.1. A Monopolistic Supplier in the Hedonic Equilibrium

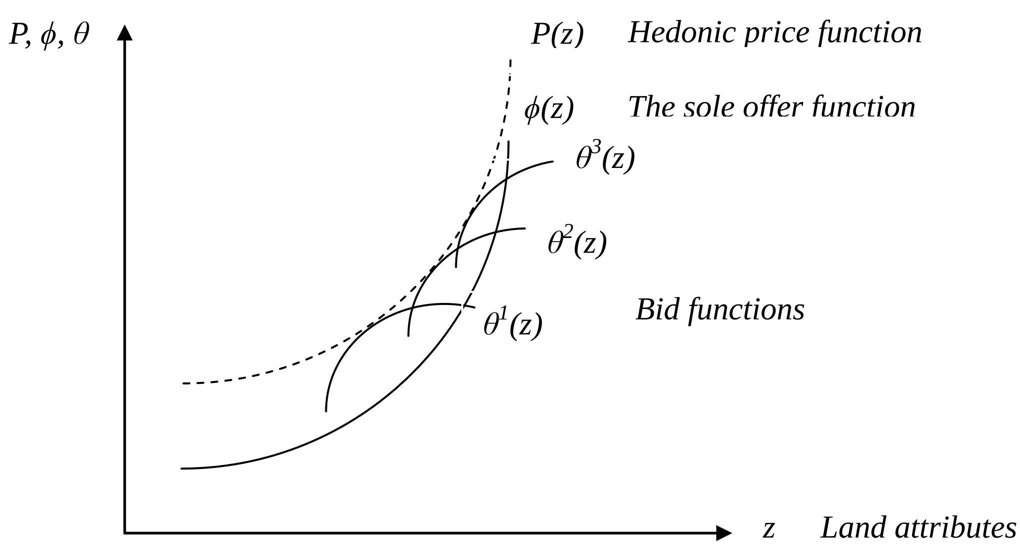

Fig. 2.2. Possible Bidders' Premium in the Hedonic Equilibrium

publicly posts the current highest bid, but allows bidders to repeatedly submit new bids over a longer period of time (e.g., two weeks). In essence, both approaches represent an open ascending-bid auction, better known as an English Auction.

In many cities in China, an authority called the "developable land reserve center" processes land for development. Land becomes available for development in two ways. First, the local government can engage in renewal of an aging city center by paying the original residents to move out, or allocating residents to alternative (generally larger and newer) housing units; old buildings are then dismantled prior to selling the land for new development. Another important source of developable land is agricultural land located in the suburban regions of a city. Although strict restrictions govern conversion of agricultural land, the cost of converting agricultural land into developable land reserve is still much lower than land located in the central portion of a city. The revenue generated from all such land sales is an important source of local public financing (at present, there is no property tax in China).[1]

3.2. The City of Chengdu

The city of Chengdu is the capital city of Sichuan Province which lies in the southwestern part of mainland China. It is situated at the western edge of the Sichuan Basin, about 1500 kilometers southwest of Beijing. With nearly 13 million official residents, Chengdu is the fourth largest city in China and serves as the most important economic, transportation and communication hubs in southwestern China. The most urbanized part of the city consists of 4 concentric ring roads, with a fifth

1 The central government and local government share the land sales revenues.

ring road under construction. It is expanding in nearly all directions via planned and *in situ* mass transportation modes (a planned subway system and an already well-developed highway system). Further, Chengdu is a standard monocentric city lying in a plain, which frees us from concerns regarding heterogeneity in hypsography. Chengdu also has very active markets in both developable land and residential housing but, as an inland city, it is subject to less speculation than the coastal cities. The natural boundary of the metropolitan area is within the fourth ring road, composed of about 600 square kilometers, though in some directions urbanization goes beyond the fourth ring area (see Fig. 2.3). Areas to the northwest, west, south, and southeast of the city center have access to high speed, low-congestion roads with easy access to the main city; they are also rich in natural open space amenities. Expansion to the west of the city center is strictly restricted due to farm land protection. Thus, most future expansion will be to the north, east, and south.

3.3. Data description

We have obtained all government land transaction records from the Bureau of Land and Resources Chengdu. The data set consists of 450 observations of land sales for residential development between January 2004 and October 2009. Parcel locations in the official sales record were manually mapped to GIS coordinates; 100 parcels either could not be located with precision or were located outside our study area and were dropped from the data set, leaving 350 land sales for residential

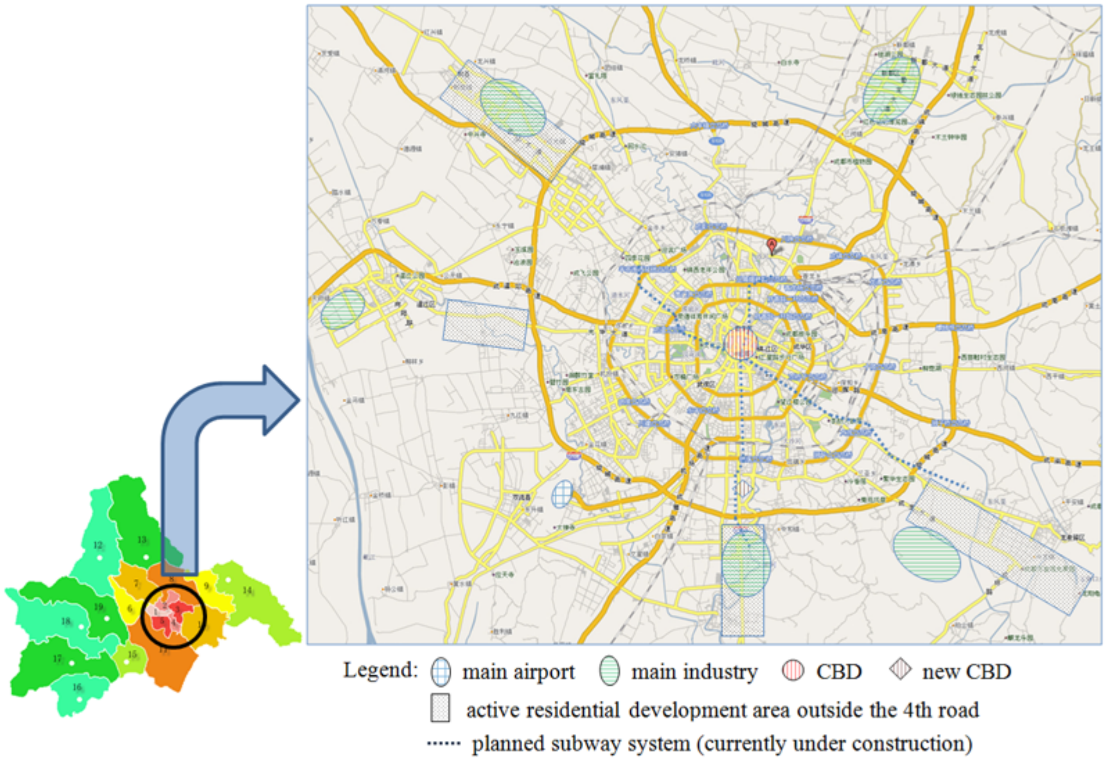

Fig. 2.3. Metropolitan Area of Chengdu

development. Fig. 2.4 shows the locations of the parcels in the data set. Some 17% of parcels were located inside the first ring road, 10% between the first and second ring roads, 34% between the second and third, 18% between the third and fourth; 21% of parcels were located outside of the fourth ring road. In addition, we also distinguish parcels by locations within the eleven administrative districts making up the study area of Chengdu city. All administrative districts are bisected by more than two ring roads, allowing us to use these two kinds of variables to capture unobserved neighborhood effects for any given parcel.

Each transaction record provides information about the transaction date, the type

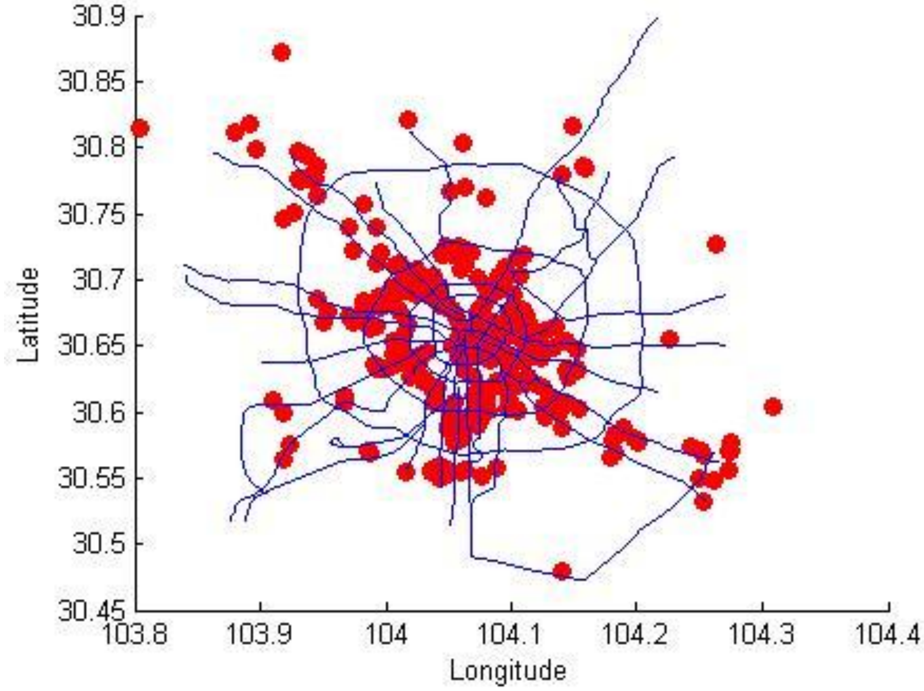

Fig. 2.4. Spatial Distribution of Land Sales

of auction governing the transaction, the area of the parcel, the per unit area transaction price as well as the asking price listed by the local government. Prices are measured as RMB ¥ per square meter.[2] Fig. 2.5 shows the spatial distribution of the unit land transaction price over the study area, both in 3-D and perpendicular views. It is easy to discern that the highest land prices lie in the center of the city, which is consistent with the prediction of a monocentric urban model. Land parcels directly south of the city center appear to have a higher price than other parcels

2 The standard "posted" price unit used for land sales in China is ¥10,000 per Chinese acre (roughly 666.667 m^2) .

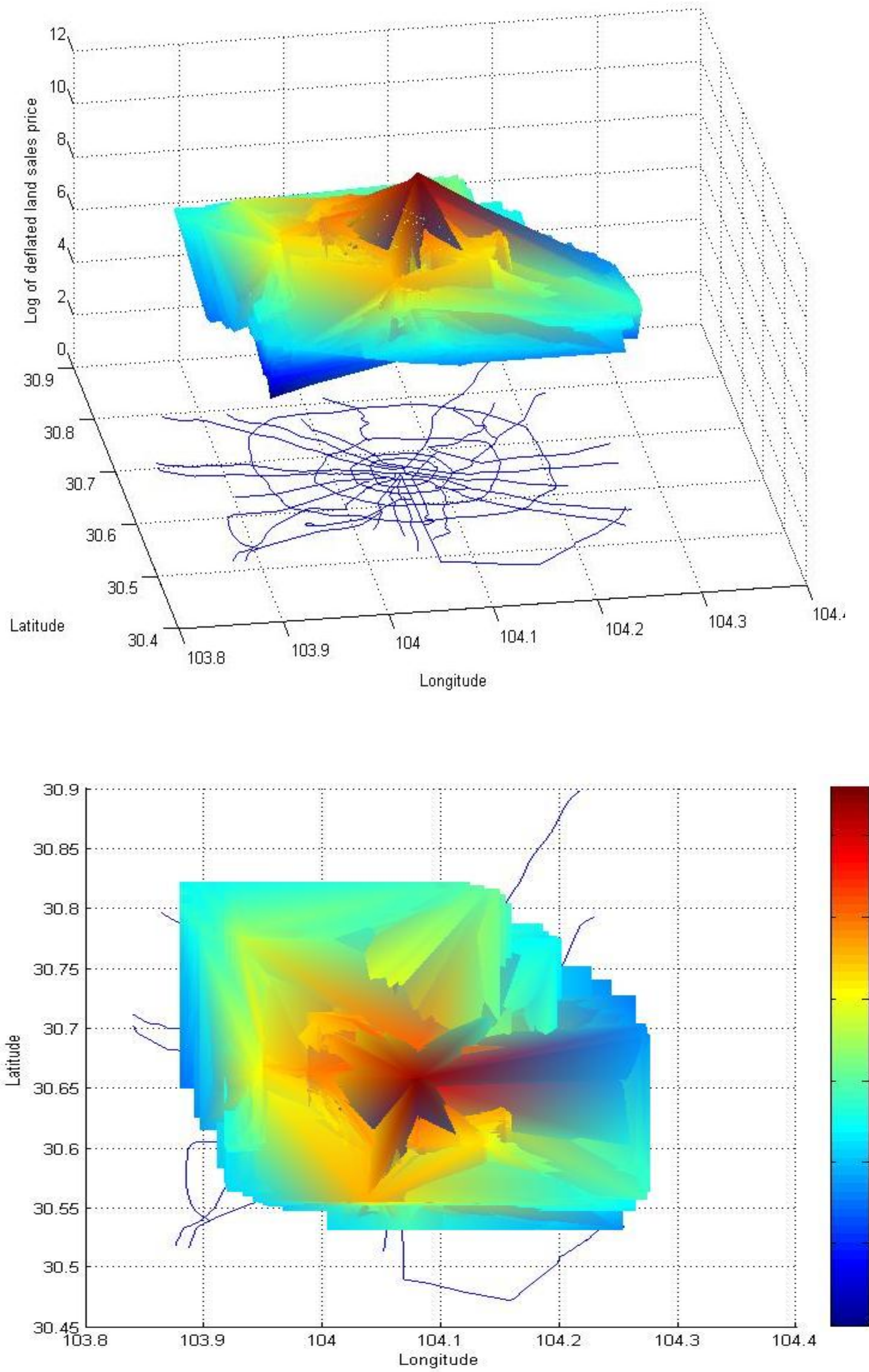

Fig. 2.5. Distribution of the Log Values of Land Sales Price over the Study Area

Table 2.1
Descriptive Statistics of Variables in Transaction Record

	Mean	Median	Std. Dev.
Asking price (￥/ m^2)	4435.430	3899.998	3732.432
Sales price (￥/ m^2)	8220.774	5999.997	9582.095
Parcel Size (m^2)	50777.200	27799.290	112558.900
Single plot (1=yes)	87.4%	--	0.332
Type 1 Auction (1=yes)	82.9%	--	0.377
Type 2 Auction (1=yes)	17.1%	--	0.377
Maximum Plot Ratio	4.089	4.000	1.607
Maximum Structural Ratio	0.325	0.300	0.085
Minimum Green Ratio	0.264	0.250	0.050

located with the same distance from the city center, but that is likely because a future central business district is currently under construction between the 3rd ring road and the 4th ring road in the south.

Land offered for sale by the government is frequently accompanied by detailed development restrictions. For example, the density of a parcel is restricted by maximum values for Plot Ratio, the ratio of total floor area (also known as construction area) to the land parcel area; the Structural Density Ratio, the ratio of the total base area of the building to the land parcel area. Structural Density essentially restricts the footprint of a building, whereas the Plot Ratio limits the overall area of a multistory building. Finally, another important development restriction is the Green Ratio, the minimum ratio of the open space area to the land parcel area. The statistics for these measures are reported in Table 2.1.

We also include five sources of open space amenities and local infrastructure that

Table 2.2
Descriptive Statistics of Variables in Proximity and Aggregate

	Mean	Median	Std. Dev.
Proximity			
Park Proximity (m)	1834.374	1260.185	1672.456
Hospital Proximity (m)	2015.939	1172.571	2318.161
Subway Station Proximity (m)	3110.657	1778.537	3523.823
River Proximity (m)	1591.718	1036.565	1673.176
Road Proximity (m)	437.529	261.535	758.167
Aggregate			
Park Level	0.115	0.121	0.039
Hospital Level	0.111	0.095	0.064
Subway Station Level	0.006	0.006	0.003
River Level	0.205	0.211	0.048
Road Level	8.908	9.199	2.314

might affect the value of land for residential development. The statistics for these variables are reported in Table 2.2. Urban amenities may include public parks, or a view of one of the many rivers flowing through Chengdu. Infrastructure that might be important to development decisions include accessibility to highways and the major roads network in the study area, as well as subway stations planned for the near future, or hospitals.[3] We capture these influences using two different measures: for some variables, such as a view of the river or distance to the nearest subway station, a simple proximity measure (distance) may be appropriate. For other variables such as accessibility to public parks or hospitals, a simple proximity variable might not be enough to capture the major value associated with accessibility. Instead, an

3 We include only publicly owned hospitals in Chengdu.

aggregate variable that captures the scale of amenities or infrastructure (the number of hectares of a park or beds in a hospital) may prove to be a better measure. To some extent, the precise measurement being used—proximity or aggregate—is an empirical matter, so we have calculated both for use in the analysis. Measures of *proximity* simply capture the shortest distance to the amenity or infrastructure, measured in meters using the Haversine Formula.[4] All proximity values are logged to take care of the scale issue (also see for example, Mahan et al., 2000). For *aggregate* measures we use a weighting formula that "discounts" amenities or infrastructure located further away from the parcel. For example, our aggregate measure of K public parks associated with a land parcel located at latitude u and longitude v is,

$$a(u,v) = \sum_{k=1}^{K} a_k/z_k \qquad (2.1)$$

where $a(u,v)$ measures the public park aggregate, a_k is the size of the k^{th} park in square meters, and z_k is the distance from land parcel to the k^{th} park. In addition to public parks, this calculation was also completed for hospitals (a_k = beds in the k^{th} hospital), subway stations (a_k = 1 for each station), river locations (a_k = 1 for 1500 river locations) and major roads (a_k = 1 for 70,826 road locations). The aggregate measures for subway stations, river locations, and roads are akin to the method used

4 The Haversine formula calculates the distance between any two points on a sphere. Haversine distance is usually obtained in the following steps: R = earth's radius (mean= 6,371km), $\Delta lat = lat_2 - lat_1$, $\Delta long = long_2 - long_1$, $a = sin^2(\Delta lat/2) + cos(lat_1) \times cos(lat_2) \times sin^2(\Delta long/2)$, $c = 2 \times arcsin\{min[1,sqrt(a)]\}$, $d = R \times c$. All angles are measured in radians.

by Gibbons and Machin (2003), where we capture not just the positive amenity of accessibility but also any disamenities that might be associated with crime (subway stops) or noise (roads). That is, high values of the aggregate road or subway measures may either positively or negatively affect parcel values, whereas high values of the aggregate river measure may be associated with the amenity of being surrounded on many sides by water.

4. English Auctions

English auctions are known to have an incentive incompatibility problem in that participants, including the winner, need not reveal their true valuations according to the auction mechanism.[5] In an auction setting, the market involves competition only on one side: a single seller versus several potential buyers. For the seller, the situation is relatively simple. As the seller announces an asking price, its true valuation can be derived from Riley and Samuelson's (1981) formula based on its asking price as well as the distribution of the buyer's valuation.[6] The situation is more complicated for bidders. An English auction is equivalent to Vickrey's second price sealed auction in the sense that the highest bidder (presumed to be the bidder with the highest true valuation) wins. However, in Vickrey's second price sealed

5 In some auction studies, a player's reservation price (or reservation value) denotes its true valuation; while in others, they are not the same. To avoid possible confusion, we do not use the term "reservation price" in this study.
6 In some auction studies, the asking price is referred to as the "reserve price."

auction, the winner's valuation is known and the winner only pays the second-highest valuation as the rule requires. Although in an English auction the winning bid asymptotically approaches the second-highest valuation, the winner's true valuation remains unobservable.[7] We explore some details in English auction with a focus on the market participants' true valuations below.

Riley and Samuelson (1981) present a method to derive the optimal asking price of the seller in an English auction. Their approach is implemented in three steps. In the first step, Riley and Samuelson derive the expected revenue of the seller. They start their derivation from the buyer, and define the buyer's expected gain as the product of true valuation, v, and probability of winning, minus the expected payment. For buyer i, its non-cooperative equilibrium bid Θ_i, is a function of true valuation v_i, hence $\Theta_i = \Theta(v_i)$. Consider a particular potential buyer, denoted by "buyer 1," who bids according to $\Theta_1 = \Theta(v)$. As Milgrom and Weber (1982) have shown that, when there are least two players to bid in an English auction, the dominant bidding strategy is $\Theta(v) = v$.[8] Assuming that there are n_p potential bidders (players) in the auction, buyer 1 wins only when all other $n_p - 1$ buyers bid less than $\Theta(v)$. Let the

7 Empirically, the English auction winner pays the second-highest valuation plus the last increment in the auction, with the last increment asymptotically approaching zero.

8 Note that for the winner, even its dominant strategy is to bid according to its true valuation, the winner does not necessarily need to pay according to its true valuation. Riley and Samuelson (1981) have similar argument, and they call such bidding strategy as the "optimal strategy" of the buyers. Therefore, as a result of the auction (not strategy), $\Theta(v) = v$ holds only for the losers in the auction. This is commonly referred to as "loser tells the truth."

cumulative distribution function, $F(v)$, show the probability a buyer has a true valuation less than or equal to v. Given the independently identical distribution (*i.i.d*) assumption, buyer 1 wins with the probability of $[F(v)]^{n_p-1}$.[9] Therefore, buyer 1's expected gain in the auction is,

$$\Pi(v, v_1) = v_1 \times [F(v)]^{n_p-1} - P(v) \tag{2.2}$$

where $P(v)$ is buyer 1's expected payment.

Buyer 1's optimal choice according to the bidding strategy of $\Theta(v)$ is $v = v_1$, thus at $v = v_1$, the following first order condition must hold:

$$\frac{\partial \Pi(v, v_1)}{\partial v} = v_1 \times \frac{d[F(v)]^{n_p-1}}{dv} - \frac{\partial P(v)}{\partial v} = 0 \tag{2.3}$$

Let us now introduce the buyer's threshold valuation regarding the auction object, r, below which it is not profitable to bid.[10] Thus, the following participation constraint must hold as well:

$$\Pi(r, r) = r \times [F(r)]^{n_p-1} - P(r) = 0 \tag{2.4}$$

Therefore, for all $v_1 \geq r$, Eq. (2.3) can be rewritten as,

---

9 The event "buyer 1 wins" is equivalent to the event "all other $n_p - 1$ potential buyers fail." Note that the probability of a potential buyer, whose valuation is less than v, is $F(v)$. Then, according to the *i.i.d.* assumption, the probability of "all other $n_p - 1$ potential buyers fail" is $[F(v)]^{n_p-1}$.

10 We call r the threshold valuation by meaning that if the buyer's valuation is exactly r ($v = r$), then its expected profit is zero. Then for those buyers whose valuation is greater than r, they are anticipating some positive level of profit. However, as the buyer increases its bid in the auction, such expected profit is gradually consumed. Finally, when the buyer bids at its true valuation (i.e., the maximum amount it can bid), the expected profit becomes zero again.

$$\frac{\partial P(v_1)}{\partial v_1} = v_1 \times \frac{d[F(v_1)]^{n_p-1}}{dv_1} \tag{2.5}$$

Buyer 1's expected payment is obtained by integrating Eq. (2.5) and using Eq. (2.4) as a boundary condition,

$$P(v_1) = v_1 \times [F(v_1)]^{n_p-1} - \int_r^{v_1} [F(v)]^{n_p-1} dv \tag{2.6}$$

Now, for the seller, both v_1 and $P(v_1)$ are random variables, but with known distribution. Hence, the seller's expected revenue from buyer 1 is $E[P(v_1)]$, as follows:

$$E[P(v_1)] = \int_r^{\bar{v}} \{[v \times \frac{dF(v)}{dv} + F(v) - 1] \times [F(v)]^{n_p-1}\} dv \tag{2.7}$$

where $\bar{v}$ is the maximum value that the random variable v can take, i.e., $F(\bar{v})=1$.[11] Since the seller has no private information about the potential buyers beyond the distribution of their true valuation, the seller uses "equal treatment" regarding all n_p buyers, i.e., every buyer might be buyer 1. Therefore, the seller's expected revenue from buyer 1 is,

$$n_p \times E[P(v_1)] = n_p \times \int_r^{\bar{v}} \{[v \times \frac{dF(v)}{dv} + F(v) - 1] \times [F(v)]^{n_p-1}\} dv \tag{2.8}$$

The second step that Riley and Samuelson (1981) implement is to derive the buyer's equilibrium bidding strategy. Assume that the seller announces an asking price, Θ_0, which is the minimum amount that the seller would accept in the auction. Obviously, only those potential buyers who have true valuation $v > \Theta_0$ would

11 Note that $\bar{v}$ is the hypothetical boundary of the distribution $F(v)$, which predicts the event that "every buyer fails." In another word, there would be no winner at $v = \bar{v}$.

participate in the auction. From the buyer's view point, the expected payment is hence,

$$P(v) = Prob\ \{\text{the buyer is the winner}\} \times \Theta(v) \tag{2.9}$$

Solving $\Theta(v)$ from Eq. (2.9) yields the buyer's equilibrium bidding strategy.

The third step that Riley and Samuelson (1981) implement is to derive the seller's optimal asking price. In Eq. (2.8), we do not consider the case that the auction fails. When the true valuations of all buyers are less than r, then no buyers will participate in the auction. The probability of such case is $[F(r)]^{n_p}$. Then, the seller would have the "gain" of its own true valuation, v_0. Thus, we could construct the seller's "total" expected return, TR, as follows:

$$E[TR] = v_0 \times [F(r)]^{n_p} + n_p \times \int_r^{\bar{v}} \{[v \times \frac{dF(v)}{dv} + F(v) - 1] \times [F(v)]^{n_p - 1}\} dv \tag{2.10}$$

Differentiating Eq. (2.10) with respect to r, we obtain the optimal value of the asking price,

$$n_p \times [v_0 \times \frac{dF(r)}{dr} - r \times \frac{dF(r)}{dr} - F(r) + 1] \times [F(r)]^{n_p - 1} = 0 \tag{2.11}$$

Rearranging Eq. (2.11), we have:

$$v_0 = r - [1 - F(r)] / f(r) \tag{2.12}$$

In Eq. (2.12), $F(r)$ is the cumulative distribution function (CDF), and $f(r)$ is the probability density function (PDF). Note that the number of potential buyers has been eliminated. Therefore, to solve for the seller's true valuation v_0, we only need information about the distribution of the buyer's valuation as well as the asking price announced by the seller.

The true valuation held by the winner is a bit more complicated to obtain. Based on Riley and Samuelson's (1981) study, we use three steps to derive the winner's true valuation. The first step is to link the true valuation of the winner and the second-highest bidder. We denote the true valuations of the winner and the second-highest bidder by v_1 and v_2, respectively. Since Eq. (2.6) holds for every potential buyer in the auction, we have:

$$P(v_2) = v_2 \times [F(v_2)]^{n_p-1} - \int_r^{v_2} [F(v)]^{n_p-1} dv \qquad (2.13)$$

Similarly to Eq. (2.9), we can write the second-highest bidder's expected payment as,

$$P(v_2) = Prob\ \{\text{the buyer is the second-highest bidder}\} \times \Theta(v_2) \qquad (2.14)$$

Then, what is the probability of a buyer being the second-highest bidder? We now divide all potential buyers into three groups: the winner, the second-highest bidder, and other buyers. All other buyers have their true valuations less than v_2, with probability $[F(v_2)]^{n_p-2}$. In addition, the winner wins only when the second-highest bidder's true valuation is less than v_1. This is given by probability $F(v_1)$. Therefore, the total probability can be expressed as follows:

$$Prob\ \{\text{the buyer is the second-highest bidder}\} = [F(v_2)]^{n_p-2} \times F(v_1) \qquad (2.15)$$

Combining Eqs. (2.13) to (2.15), we have:

$$v_2 \times [F(v_2)]^{n_p-1} - \int_r^{v_2} [F(v)]^{n_p-1} dv = [F(v_2)]^{n_p-2} \times F(v_1) \times \Theta(v_2) \qquad (2.16)$$

Since "loser tells the truth," we have: $\Theta(v_2) = v_2$. Therefore, Eq. (2.16) can be rewritten as:

$$v_2 \times [F(v_2)]^{n_p-1} - \int_r^{v_2} [F(v)]^{n_p-1} dv = [F(v_2)]^{n_p-2} \times F(v_1) \times v_2 \qquad (2.17)$$

The term $\int_r^{v_2} [F(v)]^{n_p-1} dv$ cannot be directly integrated. However, according to the Fundamental theorem of calculus, we have:

$$\frac{\partial}{\partial v_2} \{ \int_r^{v_2} [F(v)]^{n_p-1} dv \} = [F(v_2)]^{n_p-1} \qquad (2.18)$$

Thus, we differentiate both sides of Eq. (2.17) with respect to v_2, after rearrangement, we obtain the link between true valuations of the winner and the second-highest bidder, as follows:

$$F(v_2) \times F(v_1) + v_2 \times (n_p - 2) \times f(v_2) \times F(v_1) = v_2 \times (n_p - 1) \times F(v_2) \times f(v_2) \qquad (2.19)$$

Eq. (2.19) in fact shows the probability relationship between v_1 and v_2. The English auction winner pays the second-highest valuation plus the last increment in the auction. However, the increment in the auction is usually very small. Hence, in an English auction the winning bid asymptotically approaches the second-highest valuation. Denoting the actual sales price (winning bid) by s, we have $v_2 \approx s$. After obtaining the second-highest valuation v_2, the number of potential buyers n_p is yet unknown. Therefore, our second step to uncover v_1 is to find n_p.

As Paarsch (1997) has pointed out, a measure of potential competition in the auction is notoriously difficult, and often impossible. With the knowledge of the "actual bidders," whose true valuations are no less than the asking price proposed by the seller, Paarsch uses a conditional relationship to map out the potential competition upon the number of actual bidders. However, we do not have such information about the actual bidders. Recall that in Eq. (2.8), we have presented the expected revenue to the seller when the auction is successful. In real world, the seller gets the actual sales price (winning bid) as the result of a successful auction. Therefore, we have:

$$n_p \times \int_r^{\bar{v}} \{[v \times \frac{dF(v)}{dv} + F(v) - 1] \times [F(v)]^{n_p - 1}\} dv = s \qquad (2.20)$$

Solving n_p from Eq. (2.20),[12] we obtain a measure of potential competition (note that $\bar{v}$ is solved from $F(\bar{v})=1$).

The third step we need to reveal v_1 is to derive the distribution of the buyer's valuation, *F(v)* and *f(v).* Paarsch proposes a method to use the bonus bid (auction premium) to empirically estimate the distribution of the buyer's valuation. Paarsch defines the bonus bid *b*, as:

$$b = s - r \geqslant 0 \qquad (2.21)$$

where *s* is the actual sales price (winning bid) and *r* is the seller's asking price. Obviously, the bonus bid, *b*, is a variable with a non-negative value. Paarsch has proposed a conditional maximum likelihood estimator to estimate the distribution of the buyer's valuation based on the number of the actual bidders in the auction. In our study, we follow Paarsch's basic idea to derive the distribution of the auction

12 Note that the integrand $[v \times \frac{dF(v)}{dv} + F(v) - 1] \times [F(v)]^{n_p - 1}$ is highly non-linear, which makes it impossible to directly conduct the integration. Hence, we conduct the first-order Taylor expansion to linearize the integrand before we do the integration.

premium and, thus, an estimate of the winner's true valuation v_1. The exact method is presented in a later section of this paper.[13]

5. Empirical Models

Before presenting our models it is necessary to address a number of empirical issues. First, a common econometric problem in hedonic modeling is that the data are related to one another in a spatially heterogeneous manner. Anselin (1988) uses "spatial dependence" or "spatial correlation" to denote the case in which the value observed in one location depends on the values at neighboring locations. The spatial correlation problem can be addressed using either a spatial-lag model or spatial-error model, the two most common spatial econometric models (each with many variants). In our study, we only focus on the Spatial Autoregressive Model (SAR) and the Spatial Error Model (SEM). The form of the SAR model is,

$$y = \boldsymbol{\rho} \times W \times y + X \times \boldsymbol{\beta} + e \tag{2.22}$$

whereas the functional form of the SEM is given by,

$$y = X \times \boldsymbol{\beta} + u,\ u = \boldsymbol{\lambda} \times W \times u + e \tag{2.23}$$

In the Eqs. (2.22) and (2.23), X and y are standard explanatory and dependent variables. W is referred to as the spatial weight matrix; $\boldsymbol{\rho}$ and $\boldsymbol{\lambda}$ are the spatial lag coefficients in both SAR and SEM, respectively. The disturbance term e is assumed to

13 For more related studies, see Cremer and McLean (1988), Levin and Smith (1994), Levin and Smith (1996), and McAfee and Reny (1992).

be a Normal distribution, $N(0, \sigma^2)$.

For the SAR, $\boldsymbol{\rho}$ is a coefficient on the spatially lagged dependent variable, $W \times y$. To show the OLS properties of SAR, we transform Eq. (2.22) as follows:

$$y = (I - \boldsymbol{\rho} \times W)^{-1} \times X \times \boldsymbol{\beta} + (I - \boldsymbol{\rho} \times W)^{-1} \times e \tag{2.22a}$$

Therefore, the OLS estimator for $\boldsymbol{\beta}$ is,

$$\hat{\beta} = (X_L' \times X_L)^{-1} \times X_L' \times y \tag{2.22b}$$

where, $X_L = (I - \boldsymbol{\rho} \times W)^{-1} \times X$. Substituting Eq. (2.22a) into Eq. (2.22b) and expand all the terms, we have:

$$\hat{\beta} = (X_L' \times X_L)^{-1} \times X_L' \times X_L \times \boldsymbol{\beta} + (X_L' \times X_L)^{-1} \times X_L' \times (I - \boldsymbol{\rho} \times W)^{-1} \times e \tag{2.22c}$$

By inspection, from Eq. (2.22c) we have: $E[\hat{\beta}] = \boldsymbol{\beta}$, which means that the OLS estimates of $\boldsymbol{\beta}$ for the SAR is still unbiased. However, Anselin (1988) has shown that the OLS estimate for ρ is biased. To show this, Anselin (1988) proposes a simple model, which he calls "The first-order spatial AR model," as follows:

$$y = \boldsymbol{\rho} \times W \times y + e \tag{2.22d}$$

The estimator of $\boldsymbol{\rho}$ is hence,

$$\hat{\rho} = (y_L' \times y_L)^{-1} \times y_L' \times y = \boldsymbol{\rho} + (y_L' \times y_L)^{-1} \times y_L' \times e \tag{2.22e}$$

where $y_L = W \times y$. According to Anselin's explanation, $W \times y$ is not fixed in repeated sampling (which is the traditional requirement for the explanatory variables), since the observations are generated by a spatial process. Hence, we cannot pass the expectation operator over the term $(y_L' \times y_L)^{-1} \times y_L'$. Therefore, we know that $E[\hat{\rho}]$

$\neq \boldsymbol{\rho}$, and the estimator of $\boldsymbol{\rho}$ is biased. In addition, Anselin (1988) also proposes that the probability limit (*plim*) of the term $y_L' \times e$, which can be expressed as,

$$plim \ \frac{1}{n} \times (y_L' \times e) = plim \ \frac{1}{n} \times \{e' \times [(I - \boldsymbol{\rho} \times W)^{-1}]' \times W' \times e\} \qquad (2.22f)$$

will not equal zero for all non-trivial case of $\boldsymbol{\rho} \neq 0$. Therefore, the estimator of $\boldsymbol{\rho}$ is inconsistent.

A more interesting feature than the inconsistency of $\boldsymbol{\rho}$ is the change of $\boldsymbol{\beta}$'s variance - covariance matrix. By inspection of Eq. (2.22c), we can see that the variance - covariance of $\hat{\beta}$ depends on the term $(X_L' \times X_L)^{-1} \times X_L' \times (I - \boldsymbol{\rho} \times W)^{-1} \times e$. Thus, we have:

$$\begin{aligned} \mathrm{Var}[\hat{\beta}|X] = \sigma^2 \times (X_L' \times X_L)^{-1} \times X_L' \times (I - \boldsymbol{\rho} \times W)^{-1} \times [(I - \boldsymbol{\rho} \times W)^{-1}]' \\ \times X_L \times [(X_L' \times X_L)^{-1}]' \qquad (2.22g) \end{aligned}$$

Apparently, only in the trivial case of $\boldsymbol{\rho} = 0$, can $\mathrm{Var}[\hat{\beta}|X]$ be reduced to $\sigma^2 \times (X' \times X)^{-1}$. Therefore, $\mathrm{Var}[\hat{\beta}|X]$ is not consistent. As a result of this inefficiency issue, the *t* statistics of $\hat{\beta}$ will be underestimated.

In regard to the SEM, from Eq. (2.23) the OLS estimator for $\boldsymbol{\beta}$ is,

$$\begin{aligned} \hat{\beta} &= (X' \times X)^{-1} \times X' \times y = (X' \times X)^{-1} \times X' \times [X \times \boldsymbol{\beta} + (I - \boldsymbol{\lambda} \times W)^{-1} \times e] \\ &= (X' \times X)^{-1} \times X' \times X \times \boldsymbol{\beta} + (X' \times X)^{-1} \times X' \times (I - \boldsymbol{\lambda} \times W)^{-1} \times e \\ &= \boldsymbol{\beta} + (X' \times X)^{-1} \times X' \times (I - \boldsymbol{\lambda} \times W)^{-1} \times e \qquad (2.23a) \end{aligned}$$

Since the term $(X' \times X)^{-1} \times X' \times (I - \boldsymbol{\lambda} \times W)^{-1} \times e$, when taking expectation, would be zero, we have: $E[\hat{\beta}] = \boldsymbol{\beta}$, which means that the OLS estimates of $\boldsymbol{\beta}$ for the SEM is still unbiased. However, the probability limit of the term $(X' \times X)^{-1} \times X' \times (I - \boldsymbol{\lambda} \times$

$W)^{-1} \times e$,

$$plim\ \frac{1}{n} \times \{(X' \times X)^{-1} \times X' \times (I - \lambda \times W)^{-1} \times e\}$$

$$= plim\ \frac{1}{n} \times [(X' \times X)^{-1}] \times plim\ \frac{1}{n} \times \{X' \times (I - \lambda \times W)^{-1} \times e\} \qquad (2.23b)$$

will not equal zero for all non-trivial case of $\lambda \neq 0$. Therefore, the estimator of λ is inconsistent. Similar to Eq. (2.22g), we have:

$$\mathrm{Var}[\hat{\beta}|X] = \sigma^2 \times (X' \times X)^{-1} \times X' \times (I - \lambda \times W)^{-1} \times [(I - \lambda \times W)^{-1}]'$$

$$\times X \times [(X' \times X)^{-1}]' \qquad (2.23c)$$

Again, only in the trivial case of $\lambda = 0$, can $\mathrm{Var}[\hat{\beta}|X]$ be reduced to $\sigma^2 \times (X' \times X)^{-1}$. Therefore, $\mathrm{Var}[\hat{\beta}|X]$ is not consistent, and the t statistics of $\hat{\beta}$ will be underestimated.

While the generalized method of moments (GMM) is sometimes used to estimate spatial models, the most popular way is to use maximum likelihood estimation (MLE). In our study, we only present the results of MLE for the spatial models.

When the parcels in a hedonic data set are not contiguous, the spatial weight matrix is generally formed with element i, j as the inverse distance between parcels i and j (the elements in each row are normalized such that their summation equals one). Generally speaking, the combination of spatial techniques with hedonic pricing models would increase the R^2 as well as the significance of estimated coefficients in the regression (for example, see Kim et al., 2003).

Another issue is how to deal with the passage of time. Our data run from 2004 through 2009, a time of fluctuating land prices in China. In our study, we consider three approaches: deflation only by a HPI, use of only monthly time dummy variables, as well as a mix of these two approaches. First, we deflate both the asking price and winning bid price by a local monthly housing price index.[14] Our second approach is to include a monthly time trend variable, starting with January 2004 equal to one and ending with October 2009 equal to 70. Our third approach is to use a mix of the previous two approaches. We present and compare the estimation results using each of these approaches later in the paper.

5.1. From Auction Premium to the Land Seller's True Valuation

Traditional hedonic theory posits that the hedonic equilibrium arises from the interactions of sellers' offer functions and buyers' bid functions. Identification problems usually prevent one from estimating either the offer function or the bid function of market participants. In our case, though, there is only one supplier in the land wholesale market offering land in an English auction; we do not have a set of offer curves coming from different sellers. Thus, the posted asking prices of the monopoly supplier for different plots of land can be used to map out the offer curve as the characteristics of these plots differ.

We start our analysis by following Paarsch's (1997) approach and calculate the auction premium (or bonus bid), *b*, the difference between the actual sales price and

---

14 This housing price index of Chengdu is reported monthly by an authority called the National Development and Reform Commission.

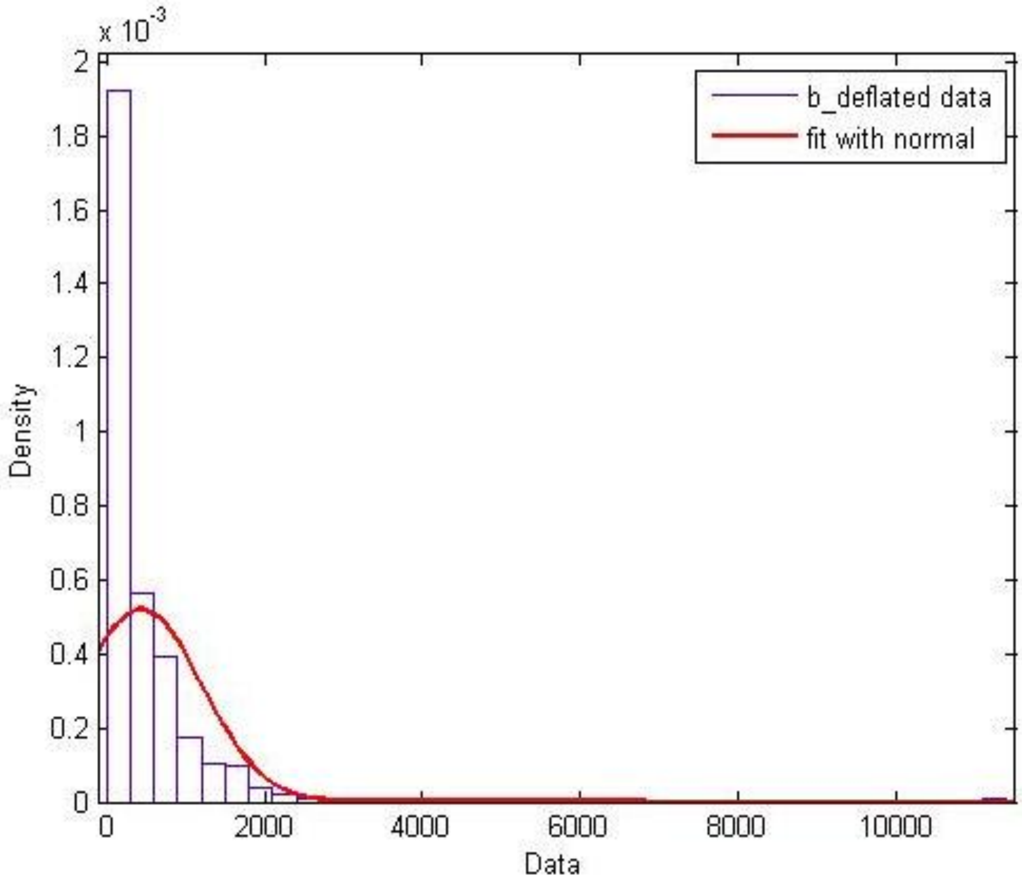

Fig. 2.6. Empirical Distribution of the Deflated Auction Premium, *b*

the asking price. The empirical distribution of deflated *b* is shown in Fig. 2.6 (where the large spike at the left includes both zero and many small non-zero values). There one may note that the empirical distribution of the auction premium follows a left-censored Normal distribution, suggesting the use of the Tobit model for its estimation. Note that, of our 350 observations of the auction premium, some 57 are equal to zero.

Paarsch's (1997) method begins with the relationship between the sales price *s*, the asking price *r*, and the bid premium *b*, as Eq. (2.21) has shown. We let *v* be the per unit raw profit of housing development net of expenditure on the land purchase. In addition, we assume that $H(L)$ is the quantity of housing arising from parcel development, P_H is the per unit housing price, and $C(L)$ is the cost of development, where L is the quantity of developable land as an input. Then the profit associated with the land input is,

$$v \times L = P_H \times H(L) - C(L) \tag{2.24}$$

Now we introduce the expenditure for land purchase, so the profit of the development, Π, is,

$$\Pi = P_H \times H(L) - C(L) - s \times L = P_H \times H(L) - C(L) - (r + b) \times L \tag{2.25}$$

Setting Π= Π*, where Π* is the desired profit level, we have:

$$\begin{aligned} b &= [P_H \times H(L) - C(L) - r \times L - \Pi^*] / L = [v \times L - r \times L - \Pi^*] / L \\ &= v - r - \Pi^*/L \end{aligned} \tag{2.26}$$

When $\Pi^* = 0$, other things equal, b achieves its maximum value, the highest bid the developer would make.[15] If we were to use a Tobit model to parameterize b according to $b = \boldsymbol{\beta}X + e$, one could use the error distribution to recover the distribution of v, which is the true valuation of the land to the developer.

We do so by noting that e is assumed to be an *i.i.d.* random variable normally distributed as $N(0,\sigma^2)$. Let $\boldsymbol{\beta}_{\text{Tobit}}$ and σ_{Tobit} be the estimation results from Tobit regression of b on the explanatory variables X, so the pdf of e, $f_e(e)$, could be denoted

15 Note that $\Pi^* \geq 0$.

as $N(0, \sigma_{Tobit}{}^2)$.[16] By the inverse function of e, $e = b - \boldsymbol{\beta}_{\text{Tobit}}X$, we can derive the pdf of b by the simple probability transformation, $f_b(b) = f_e(b - \boldsymbol{\beta}_{Tobit}X)$.[17] Noting that $b = v - u$, we could get the pdf of v by an equivalent probability transformation in a similar manner,

$$f_v(v) = f_b(v - u) = f_e(v - u - \boldsymbol{\beta}_{Tobit}X) \tag{2.27}$$

Once we have the pdf of the buyer's true valuation $f_v(v)$, we can obtain the corresponding cdf, $F_v(v)$, by integrating $f_v(v)$. Then, along with the data of the asking price proposed by the seller, we can calculate the seller's true valuation v_0 from Eq. (2.12). As soon as we have the information of v_0, we can conduct the hedonic estimation for the seller. All models were estimated using OLS, SAR, and SEM techniques. We do not go details of the tests for spatial correlation, but three out of five spatial tests suggest that there is strong spatial dependence / correlation for the seller's true valuation, and hence we report results from our SEM model.[18,19]

Our best results—on the basis on expected coefficient signs, the spatial correlation tests, and best fit—were obtained with semi-log specification using a

---

16 The estimation results of the Tobit model are listed in Tables 2.3, 2.4,and 2.5 for the cases using only HPI deflation, only monthly dummy, and a mix of the two, respectively. Since the Tobit model estimation is just an intermediate step in this section, we do not discuss its results in detail.

17 Note that $f_b(b) = f_e(b - \boldsymbol{\beta}_{Tobit}X) \times |\mathrm{d}e / \mathrm{d}b|$, and $\mathrm{d}e / \mathrm{d}b = 1$.

18 We used test statistics for Moran's I-test, a likelihood ratio test, a Wald test, and a Lagrange multiplier test for spatial correlation in the residuals, and a Lagrange multiplier test for correlation in the SAR residuals. See Anselin (1988) and LeSage (1999) for details.

19 In our study, SEM does a better job than SAR estimation in terms of higher t statistic values, R^2 value, and log-likelihood value, as well as "correct" signs of the estimated coefficients which are consistent with our expectation.

Table 2.3
Tobit Estimation of Auction Premium, *b* (￥/ m^2), Using HPI Deflation

Variable	Coefficient	t-statistic	p-value
Intercept	12369.698	0.811	0.418
Development Restrictions			
Single plot (1=yes)	1498.475	1.555	0.121
Maximum Plot Ratio	53.184	0.230	0.818
Maximum Structural Ratio	-4796.245	-1.124	0.262
Minimum Green Ratio	-43140.535	-5.516	0.000
Public Amenities and Infrastructure			
ln(Subway Station Proximity)	9.755	0.015	0.988
ln(River Proximity)	189.119	0.545	0.586
ln(Park Proximity)	-758.981	-1.261	0.208
Hospital Aggregate	-316.350	-0.020	0.984
Road Aggregate	115.937	0.228	0.820
District and Ring Road Dummy Variables			
Jin Niu	-1136.545	-0.981	0.327
QingYang	-607.044	-0.528	0.598
Cheng Hua	-1734.941	-1.640	0.102
Wu Hou	-881.698	-0.750	0.454
Gao Xin South	-1294.057	-1.020	0.309
Gao Xin West	-1062.407	-0.388	0.698
Long Quan	-3145.153	-1.637	0.103
Pi County	-2813.957	-1.390	0.166
Shuang Liu County	-2265.588	-1.061	0.289
Xin Du	-1604.958	-0.745	0.457
Within 1st Ring	1104.739	0.560	0.577
Between 1st and 2nd Ring	-788.132	-0.585	0.559
Between 3rd and 4th Ring	309.609	0.296	0.767
Outside 4th Ring	988.627	0.555	0.579
Other Variables			
Type2Auction (1=yes)	-2194.550	-2.760	0.006
ln(Parcel Size)	576.185	1.762	0.079

Dependent variable: Deflated Auction Premium

Table 2.4
Tobit Estimation of Auction Premium, *b* (¥/ m^2), Using Monthly Time Dummy

Variable	Coefficient	t-statistic	p-value
Intercept	8252.940	0.766	0.444
Development Restrictions			
Single plot (1=yes)	1972.591	1.677	0.094
Maximum Plot Ratio	-186.425	-0.630	0.529
Maximum Structural Ratio	-2083.821	-0.386	0.701
Minimum Green Ratio	-45830.917	-4.272	0.000
Public Amenities and Infrastructure			
ln(Subway Station Proximity)	155.813	0.267	0.790
ln(River Proximity)	275.323	0.666	0.506
ln(Park Proximity)	-1044.481	-1.539	0.125
Hospital Aggregate	2996.909	0.148	0.882
Road Aggregate	270.746	0.583	0.560
District and Ring Road Dummy Variables			
Jin Niu	-1719.465	-1.187	0.236
QingYang	-1383.771	-0.920	0.358
Cheng Hua	-2079.755	-1.504	0.134
Wu Hou	-1674.506	-1.150	0.251
Gao Xin South	-1663.596	-1.009	0.314
Gao Xin West	-2504.278	-0.723	0.470
Long Quan	-2884.502	-1.173	0.242
Pi County	-3417.104	-1.376	0.170
Shuang Liu County	-1704.494	-0.607	0.545
Xin Du	-4662.520	-1.613	0.108
Within 1st Ring	2235.057	0.878	0.381
Between 1st and 2nd Ring	-911.437	-0.546	0.586
Between 3rd and 4th Ring	229.441	0.171	0.864
Outside 4th Ring	718.931	0.314	0.754
Other Variables			
Type2Auction (1=yes)	-2956.123	-2.836	0.005
ln(Parcel Size)	752.266	1.964	0.050
Time Trend	75.201	2.982	0.003

Dependent variable: Auction Premium

Table 2.5
Tobit Estimation of Auction Premium, *b* (￥/ m^2), Using Both HPI Deflation and Monthly Dummy

Variable	Coefficient	t-statistic	p-value
Intercept	6568.799	0.798	0.425
Development Restrictions			
Single plot (1=yes)	1550.168	1.744	0.082
Maximum Plot Ratio	-93.615	-0.419	0.675
Maximum Structural Ratio	-1436.690	-0.351	0.726
Minimum Green Ratio	-34189.700	-4.224	0.000
Public Amenities and Infrastructure			
ln(Subway Station Proximity)	103.014	0.232	0.817
ln(River Proximity)	182.405	0.584	0.560
ln(Park Proximity)	-800.054	-1.562	0.119
Hospital Aggregate	2137.599	0.140	0.889
Road Aggregate	211.512	0.601	0.548
District and Ring Road Dummy Variables			
Jin Niu	-1354.730	-1.238	0.216
QingYang	-1049.460	-0.925	0.356
Cheng Hua	-1639.060	-1.570	0.117
Wu Hou	-1275.920	-1.161	0.246
Gao Xin South	-1156.830	-0.929	0.353
Gao Xin West	-1833.970	-0.701	0.484
Long Quan	-2239.700	-1.207	0.228
Pi County	-2575.810	-1.373	0.171
Shuang Liu County	-1330.000	-0.627	0.531
Xin Du	-3307.300	-1.517	0.130
Within 1^{st} Ring	1728.213	0.899	0.369
Between 1^{st} and 2^{nd} Ring	-713.114	-0.566	0.572
Between 3^{rd} and 4^{th} Ring	140.318	0.139	0.890
Outside 4^{th} Ring	516.191	0.299	0.765
Other Variables			
Type2Auction (1=yes)	-2376.940	-3.020	0.003
ln(Parcel Size)	531.855	1.837	0.067
Time Trend	58.695	3.093	0.002

Dependent variable: Deflated Auction Premium combination of proximity and aggregate measures for public good amenities and infrastructure. Estimation results are shown in Tables 2.6, 2.7 and 2.8 for the three cases regarding different time approaches. The one with HPI deflation (Table 2.6) has the smallest R^2 and log likelihood value. In addition, some of the coefficients' signs are not consistent with our expectation. The models using monthly time dummy (Table 2.7) and both deflation and dummy (Table 2.8) have roughly similar results. However, since some of the key variables in the mixed case have slightly larger *t*-values, and the value of log likelihood is also larger, we consider the one using a mix of deflation and dummy to be the best model specification.

Using the logarithm of the seller's derived true valuation v_0 as the dependent variable, we find that the only development restriction that the seller takes into account is the maximum Plot Ratio (total floor area relative to parcel size): as the maximum Plot Ratio increases its derived true valuation increases. The seller also notes the value of proximity to a planned subway station in that the true valuation falls as the plot gets further away. Another infrastructure measure that affects the true valuation is the aggregate measure of hospitals. That is, as the number of hospital beds, inversely weighted by distance to the hospital, increases, the seller's true valuation increases. In addition, seven of the ten district variables were statistically significant, indicating that location within the city does affect the seller's true valuation for the parcel, and the neighborhood effects exist to some degree. Relative to the baseline location between the second and third ring roads, from the

Table 2.6
Spatial Error Model of Seller's Derived True Valuation, v_0* (￥/ m^2), Using HPI Deflation

Variable	Coefficient	t-statistic	p-value
Intercept	8.8149	13.6207	0.000
Development Restrictions			
Single plot (1=yes)	-0.0170	-0.24346	0.80764
Maximum Plot Ratio	0.09246	5.2577	0.000
Maximum Structural Ratio	-1.1689	-3.6626	0.00025
Minimum Green Ratio	-2.0975	-3.41086	0.00064
Public Amenities and Infrastructure			
ln(Subway Station Proximity)	-0.10238	-2.84222	0.00448
ln(River Proximity)	-0.0111	-0.42128	0.67354
ln(Park Proximity)	0.03827	0.88762	0.374744
Hospital Aggregate	3.7053	2.91580	0.00354
Road Aggregate	-0.00088	-0.03019	0.97590
District and Ring Road Dummy Variables			
Jin Niu	-0.1072	-1.12512	0.26053
QingYang	0.11064	1.12379	0.2610
Cheng Hua	-0.1976	-2.2023	0.02763
Wu Hou	0.16196	1.72759	0.0840
Gao Xin South	-0.0587	-0.53766	0.5908
Gao Xin West	-0.26347	-1.2110	0.225873
Long Quan	-0.8123	-5.2328	0.000
Pi County	-0.73076	-4.4840	0.000
Shuang Liu County	-1.1720	-6.6398	0.000
Xin Du	-0.3698	-2.1356	0.03270
Within 1^{st} Ring	-0.09579	-0.6071	0.5437
Between 1^{st} and 2^{nd} Ring	0.13659	1.32154	0.18632
Between 3^{rd} and 4^{th} Ring	-0.0052	-0.06203	0.9505
Outside 4^{th} Ring	0.06927	0.47382	0.63562
Other Variables			
Type2Auction (1=yes)	-0.1124	-1.8569	0.0633
ln(Parcel Size)	0.05349	2.2721	0.0230
λ	0.4000	1.7166	0.0860
Adjusted R-square		0.6732	
sigma^2		0.1492	
log-likelihood		-42.97258	

Dependent variable: $\ln(v_0*)$

Table 2.7
Spatial Error Model of Seller's Derived True Valuation, v_0* (￥/ m^2), Using Monthly Time Dummy

Variable	Coefficient	t-statistic	p-value
Intercept	7.244	13.024	0.000
Development Restrictions			
Single plot (1=yes)	0.009	0.157	0.875
Maximum Plot Ratio	0.056	3.786	0.000
Maximum Structural Ratio	-0.322	-1.185	0.236
Minimum Green Ratio	0.107	0.201	0.841
Public Amenities and Infrastructure			
ln(Subway Station Proximity)	-0.075	-2.478	0.013
ln(River Proximity)	-0.009	-0.390	0.697
ln(Park Proximity)	0.032	0.874	0.382
Hospital Aggregate	4.505	4.229	0.000
Road Aggregate	0.023	0.920	0.358
District and Ring Road Dummy Variables			
Jin Niu	-0.169	-2.059	0.040
QingYang	-0.005	-0.064	0.949
Cheng Hua	-0.177	-2.314	0.021
Wu Hou	0.055	0.695	0.487
Gao Xin South	-0.030	-0.323	0.747
Gao Xin West	-0.455	-2.477	0.013
Long Quan	-0.559	-4.192	0.000
Pi County	-0.662	-4.773	0.000
Shuang Liu County	-0.913	-6.130	0.000
Xin Du	-0.851	-5.660	0.000
Within 1st Ring	0.059	0.444	0.657
Between 1st and 2nd Ring	0.162	1.883	0.060
Between 3rd and 4th Ring	-0.045	-0.637	0.524
Outside 4th Ring	-0.066	-0.534	0.593
Other Variables			
Type2Auction (1=yes)	-0.168	-3.376	0.001
ln(Parcel Size)	0.044	2.267	0.023
Time Trend	0.021	16.339	0.000
λ	0.592	3.132	0.002
Adjusted R-square		0.794	
sigma^2		0.101	
log-likelihood		25.000	

Dependent variable: ln(v_0*)

Table 2.8
Spatial Error Model of Seller's Derived True Valuation, v_0* (￥/ m^2), Using Both HPI Deflation and Monthly Dummy

Variable	Coefficient	t-statistic	p-value
Intercept	7.271	13.490	0.000
Development Restrictions			
Single plot (1=yes)	-0.000	-0.002	0.998
Maximum Plot Ratio	0.053	3.687	0.000
Maximum Structural Ratio	-0.283	-1.077	0.281
Minimum Green Ratio	0.161	0.311	0.756
Public Amenities and Infrastructure			
ln(Subway Station Proximity)	-0.081	-2.734	0.006
ln(River Proximity)	-0.010	-0.458	0.647
ln(Park Proximity)	0.032	0.920	0.357
Hospital Aggregate	4.429	4.292	0.000
Road Aggregate	0.021	0.877	0.381
District and Ring Road Dummy Variables			
Jin Niu	-0.162	-2.037	0.042
QingYang	0.002	0.030	0.976
Cheng Hua	-0.176	-2.383	0.017
Wu Hou	0.058	0.745	0.457
Gao Xin South	-0.017	-0.191	0.848
Gao Xin West	-0.448	-2.518	0.012
Long Quan	-0.555	-4.289	0.000
Pi County	-0.661	-4.910	0.000
Shuang Liu County	-0.906	-6.275	0.000
Xin Du	-0.792	-5.434	0.000
Within 1st Ring	0.047	0.366	0.715
Between 1st and 2nd Ring	0.155	1.858	0.063
Between 3rd and 4th Ring	-0.052	-0.759	0.448
Outside 4th Ring	-0.073	-0.603	0.546
Other Variables			
Type2Auction (1=yes)	-0.152	-3.160	0.002
ln(Parcel Size)	0.043	2.312	0.021
Time Trend	0.016	12.914	0.000
λ	0.600	3.212	0.001
Adjusted R-square		0.788	
sigma^2		0.095	
log-likelihood		36.229	

Dependent variable: ln(v_0*)

seller's perspective only a location between the first and second ring roads has a premium associated with it. In addition, when the government offers land in a Type 2 auction, its true valuation falls. As the parcel size increases the seller's true valuation increases, too. Even after adjusting for the housing price index, the government's true valuation has tended to increase with time. Finally, the statistical significance of λ suggests spatial correlation in the data.

5.2. The Winner's True Valuation

Having estimated the elements of the government's offer function for developable land, it is now time to turn to the buyer's (developer's) side. As we have noted, the winning bid does not necessarily reveal the true valuation held by developer. Following the three steps to derive the winner's true valuation v_1 as described in section 4, we now have all the information we need. We then use the buyer's derived true valuation v_1 to estimate the bid function of developers. Tests for spatial correlation show that one of the five spatial tests suggests spatial dependence; we therefore use SEM estimation which performs better than SAR estimates. The estimation results appear in Tables 2.9, 2.10, and 2.11, for the cases using HPI deflation, monthly time dummy, and a mix of the two, respectively. The model with the HPI deflation performs worst, in the sense that it has the smallest R^2 and log-likelihood value, and its estimates of the five environmental and infrastructure variables are not statistically significant. The estimation results of the monthly time dummy variable model and the model with the mix of HPI deflation and the monthly dummy are roughly similar. Although the monthly time dummy variable model (Table 2.10) has the largest R^2 value, the *t*-values are less significant for some of the key variables than the mixed model. Therefore, we choose the model with the mix of HPI deflation and the monthly dummy as the best model specification (Table 2.11).

Using the logarithm of the land buyers' derived true valuation as the dependent variable, we find that development restrictions have a greater impact on buyers' valuation than those on the government's valuation. All else equal, developers value the land higher if the land parcel is a single plot. In addition, as the maximum Plot Ratio (total floor area relative to parcel size) increases, the buyers' valuations increase. As the maximum Structural Ratio (footprint area relative to parcel size) falls, buyers' valuations increase. Also, as the minimum Green Ratio increases, the value of land for development falls. Developers also value public amenities and infrastructure a little differently from the government. In contrast to the government, which appears to have to respond to hospital beds and planned subway infrastructure, developers place value on hospital beds and existing road infrastructure. The greater the aggregate service levels of healthcare and roads, the greater the value for development. There appears to be a strong correlation between how the government values various districts and the developers value land in those districts: of the seven negative and significant district variables in the government's offer function, developers had similar sign and significance for all seven districts.

Table 2.9
Spatial Error Model of Buyer's Derived True Valuation, v_1* (￥/ m^2), Using HPI Deflation

Variable	Coefficient	t-statistic	p-value
Intercept	9.477	10.698	0.000
Development Restrictions			
Single plot (1=yes)	0.116	1.217	0.224
Maximum Plot Ratio	0.132	5.497	0.000
Maximum Structural Ratio	-2.178	-5.013	0.000
Minimum Green Ratio	-4.854	-5.802	0.000
Public Amenities and Infrastructure			
ln(Subway Station Proximity)	-0.080	-1.613	0.107
ln(River Proximity)	-0.016	-0.441	0.659
ln(Park Proximity)	0.004	0.071	0.944
Hospital Aggregate	2.670	1.536	0.125
Road Aggregate	0.045	1.132	0.257
District and Ring Road Dummy Variables			
Jin Niu	-0.206	-1.564	0.118
QingYang	0.068	0.505	0.614
Cheng Hua	-0.371	-3.008	0.003
Wu Hou	0.101	0.787	0.431
Gao Xin South	-0.048	-0.318	0.750
Gao Xin West	-0.198	-0.665	0.506
Long Quan	-1.109	-5.195	0.000
Pi County	-0.758	-3.382	0.001
Shuang Liu County	-1.204	-4.984	0.000
Xin Du	-0.663	-2.791	0.005
Within 1st Ring	-0.095	-0.442	0.658
Between 1st and 2nd Ring	0.011	0.080	0.936
Between 3rd and 4th Ring	0.032	0.277	0.782
Outside 4th Ring	0.062	0.311	0.756
Other Variables			
Type2Auction (1=yes)	-0.536	-6.502	0.000
ln(Parcel Size)	0.058	1.795	0.000
λ	0.457	2.061	0.039
Adjusted R-square	0.651		
sigma^2	0.276		
log-likelihood	-150.943		

Dependent Variable: ln(v_1*)

Table 2.10
Spatial Error Model of Buyer's Derived True Valuation, v_1* (￥/ m^2), Using Monthly Time Dummy

Variable	Coefficient	t-statistic	p-value
Intercept	7.420	9.434	0.000
Development Restrictions			
Single plot (1=yes)	0.146	1.749	0.080
Maximum Plot Ratio	0.079	3.698	0.000
Maximum Structural Ratio	-1.084	-2.763	0.006
Minimum Green Ratio	-1.946	-2.508	0.012
Public Amenities and Infrastructure			
ln(Subway Station Proximity)	-0.040	-0.942	0.346
ln(River Proximity)	-0.012	-0.378	0.705
ln(Park Proximity)	-0.012	-0.227	0.820
Hospital Aggregate	3.614	2.395	0.017
Road Aggregate	0.077	2.214	0.027
District and Ring Road Dummy Variables			
Jin Niu	-0.287	-2.541	0.011
QingYang	-0.096	-0.824	0.410
Cheng Hua	-0.335	-3.157	0.002
Wu Hou	-0.048	-0.428	0.668
Gao Xin South	-0.014	-0.106	0.915
Gao Xin West	-0.485	-1.874	0.061
Long Quan	-0.811	-4.371	0.000
Pi County	-0.690	-3.576	0.000
Shuang Liu County	-0.892	-4.226	0.000
Xin Du	-1.335	-6.316	0.000
Within 1st Ring	0.112	0.594	0.553
Between 1st and 2nd Ring	0.046	0.380	0.704
Between 3rd and 4th Ring	-0.025	-0.243	0.808
Outside 4th Ring	-0.089	-0.514	0.607
Other Variables			
Type2Auction (1=yes)	-0.602	-8.346	0.000
ln(Parcel Size)	0.046	1.646	0.100
Time Trend	0.026	14.366	0.000
λ	0.372	1.562	0.118
Adjusted R-square	0.752		
sigma^2	0.211		
log-likelihood	-103.412		

Dependent Variable: ln(v_1*)

Table 2.11
Spatial Error Model of Buyer's Derived True Valuation, v_1* (￥/ m^2), Using Both HPI Deflation and Monthly Dummy

Variable	Coefficient	t-statistic	p-value
Intercept	7.439	9.760	0.000
Development Restrictions			
Single plot (1=yes)	0.150	1.856	0.063
Maximum Plot Ratio	0.077	3.708	0.000
Maximum Structural Ratio	-1.025	-2.700	0.007
Minimum Green Ratio	-1.893	-2.520	0.012
Public Amenities and Infrastructure			
ln(Subway Station Proximity)	-0.044	-1.070	0.285
ln(River Proximity)	-0.015	-0.497	0.619
ln(Park Proximity)	-0.011	-0.222	0.824
Hospital Aggregate	3.553	2.430	0.015
Road Aggregate	0.076	2.272	0.023
District and Ring Road Dummy Variables			
Jin Niu	-0.282	-2.576	0.010
QingYang	-0.089	-0.787	0.431
Cheng Hua	-0.340	-3.301	0.001
Wu Hou	-0.048	-0.446	0.656
Gao Xin South	0.002	0.013	0.989
Gao Xin West	-0.477	-1.900	0.057
Long Quan	-0.812	-4.511	0.000
Pi County	-0.693	-3.704	0.000
Shuang Liu County	-0.893	-4.363	0.000
Xin Du	-1.263	-6.164	0.000
Within 1st Ring	0.114	0.624	0.533
Between 1st and 2nd Ring	0.039	0.331	0.740
Between 3rd and 4th Ring	-0.034	-0.346	0.729
Outside 4th Ring	-0.095	-0.562	0.574
Other Variables			
Type2Auction (1=yes)	-0.596	-8.527	0.000
ln(Parcel Size)	0.045	1.678	0.093
Time Trend	0.021	11.930	0.000
λ	0.384	1.627	0.104
Adjusted R-square	0.750		
sigma^2	0.198		
log-likelihood	-92.024		

Dependent Variable: ln(v_1*)

Developers' true values for land were not significant in the dummy variables of ring roads. In addition, land offered at a Type 2 auction affects the value of land significantly in a negative manner, which demonstrates that the longer developers consider making a land transaction, the lower the bid. Finally, developers' valuations are found to be positively associated with parcel size, which reveals the fact that developers are more willing to pay for larger land parcels for property development. Similar to the model for the government's true valuation, developers' valuations for land have increased over time. Although the spatial correlation coefficient λ is only significant at 10.4% level, it shows that spatial correlation in the error term exists at least to some degree in the buyer's true valuation.

6. Concluding Remarks

We have thus far estimated the effect of open space and local public infrastructure on the value of urban land in a market that does not satisfy the usual assumptions of a competitive market structure, as well as incentive incompatibility issues for transaction participants. Our study shows that when confronting these two violations to the traditional assumptions of hedonic theory, we cannot directly apply a standard econometric model, due to the monopolistic seller and the incentive incompatibility issue in the English auction. Instead, we take advantage of these market features in two ways. First, the "asking price" of the government seller is used to derive its true valuation, so that one can estimate the offer function of the monopoly seller. On the buyer's side, following Paarsch's (1997) approach, we recover the distribution of the buyer's true valuation from a Tobit model estimation with respect to the auction premium, and then conduct a three-step approach based on Riley and Samuelson's (1981) study to implicitly solve for the winning buyer's true valuation through numerical methods. When we have estimated the true valuation of the winning buyers, the explanatory variables account for the buyers' derived true valuation fairly well, which allows us to estimate marginal values commonly reported in the literature.

In addition, these two violations to the traditional hedonic theory also generate two separate valuations on land with differentiated characteristics. We find that the seller and buyers differ in their marginal valuations of these land characteristics to some degree. While both placing a high value on local infrastructure (such as healthcare service level), the local government (i.e., the monopolistic land seller) values subway station proximity highly, but land buyers (i.e., the developers) exhibit higher values for road service level. In regard to proximity to subway stations, while the government considers it to be a significantly positive factor in determining property value, developers do not, perhaps because the subway system in Chengdu is still under construction. In addition, our results show that location relative to a park or a river does not matter to either the local government or the developers.

Regulation requirements for land development matter both to the seller and to the buyers. Notably, the maximum requirement for Plot Ratio significantly affects both the land seller and buyers' valuations in a positive manner, which suggests that the Plot Ratio is, perhaps, the most important economic regulation requirement on land development. While the maximum Structure Density Ratio and the minimum requirement of Green Land Ratio have negative impacts on the land buyers, they do not significantly affect the land seller.

While developers prefer parcels which consist of a single plot, the land seller does not. Locations within the various ring roads are not significant to the developers, however, "Between 1st and 2nd Ring" has been found significant to the land seller's valuation in the sense that the closer to the center of the city, the higher the land seller values. Unobserved neighborhood effects, as measured by district variables, have a significant impact on the land valuation for both the land seller and buyers. Since our omitted district (Jin Jiang District) includes a large part of the most commercialized downtown area in the city, generally speaking, the suburban districts are significantly less valued than those in the downtown area. For all the participants, the Type 1 auction tends to increase land valuation, but its impact is slightly larger for the developers than for the local government. Parcel size is also found to have a significantly positive impact on land valuation of all participants.

Our study is valuable to the land seller. It addresses one of the core issues in China's local public financing in that the local governments rely heavily on land sales for revenue generation, which is usually referred to as the "land financing." After learning the developer's true valuation on a particular land parcel with given attributes, the local government hence can increase the asking price and therefore generate more profit. In the future study, we can examine the use of alternative methods, such as a property tax, to replace the land sales in public financing, while keeping the welfare level in the economy stable.

References

Anderson, S. T., West, S. E., 2006. Open space, residential property values, and spatial context. Regional Science and Urban Economics 36, 773–789.

Anselin, L., 1988. Spatial econometrics: methods and models. Kluwer Academic Press. Dordrecht.

Asabere, P. K., Huffman, F. E., 2009. The relative impacts of trails and greenbelts on home price. Journal of Real Estate Finance and Economics 38, 408–419, doi: 10.1007/s11146-007-9089-8.

Bates, L. J., Santerre, R. E., 2001. The public demand for open space: the case of Connecticut communities. Journal of Urban Economics 50 (1), 97–111.

Bolitzer, B., Netusil, N. R., 2000. The impact of open spaces on property values in Portland, Oregon. Journal of Environmental Management 59, 185–193, doi: 10.1006/jema.2000.0351.

Cheshire, P. C., Sheppard, S. C., 1995. On the price of land and the value of amenities. Economica 62 (127), 247–267.

Cremer, J., McLean, R. P., 1988. Full extraction of the surplus in Bayesian and dominant strategy auctions. Econometrica 56, 1247–57.

Diewert, W. E., Hann, J. d., Hendriks, R., 2010. The decomposition of a house price index into land and structures components: a hedonic regression approach. Department of Economics, The University of British Columbia, Discussion Paper 10-01 (April).

Freeman, A. M., 1993. The measurement of environmental and resource values:

theory and methods. Resources for the Future. Washington, DC.

Geoghegan, J., 2002. The value of open spaces in residential land use. Land Use Policy 19 (1), 91–98.

Geoghegan, J., Wainger, L. A., Bockstael, N. E., 1997. Spatial landscape indices in a hedonic framework: an ecological economics analysis using GIS. Ecological Economics 23 (3), 251–264.

Gibbons, S., Machin, S., 2003. Rail access and house prices: an evaluation of the wider benefits of transport. Department of Transport Report , United Kingdom (November).

Irwin, E. G., 2002. The effects of open space on residential property values. Land Economics 78 (4), 465–480.

Kim, C. W., Phipps, T. T., Anselin, L., 2003. Measuring the benefits of air quality improvement: a spatial hedonic approach. Journal of Environmental Economics and Management 45, 24–39.

Leggett, C., Bockstael, N., 2000. Evidence of effects of water quality on residential land prices. Journal of Environmental Economics and Management 39, 121–144, doi: 10.1006/jeem.1999.1096.

LeSage, J. P., 1999. The theory and practice of spatial econometrics. Available at: www.spatial-econometrics.com/html/sbook.pdf

Levin, D., Smith, J. L., 1994. Equilibrium in auctions with entry. American Economic Review 84 (June), 585–99.

Levin, D., Smith, J. L., 1996. Optimal reservation prices in auctions. Economic Journal 106 (438) (September), 1271–1283.

Lutzenhiser, M., Netusil, N. R., 2001. The effect of open spaces on a home's sales price. Contemporary Economic Policy 19 (3), 291–298.

Mahan, B. L., Polasky, S., Adams, R. M., 2000. Valuing urban wetlands: a property price approach. Land Economics 76 (1), 100–113.

McAfee, P. R., Reny, P. J., 1992. Correlated information and mechanism design. Econometrica 60, 395–421.

Melser, D. 2005, The hedonic regression time-dummy method and the monotonicity axiom. Journal of Business and Economic Statistics 23 (4) (October), 485–92.

Milgrom, P. R., Weber, R. J. 1982. A theory of auctions and competitive bidding. Econometrica 50, 1089–1122.

Nelson, J. P., 1982. Highway noise and property values: a survey of recent evidence. Journal of Transport Economics and Policy 16 (2) (May), 117–138.

Paarsch, H. J., 1997. Deriving an estimate of the optimal reserve price: an application to British Columbian timber sales. Journal of Econometrics 78, 333–357.

Palmquist, R. B., 1989. Land as a differentiated factor of production: a hedonic model and its implications for welfare measurement. Land Economics 65 (February), 23–28.

Provencher, B., Sarakinos, H., Meyer, T., 2008. Does small dam removal affect local property values? an empirical analysis. Contemporary Economic Policy 26 (2) (April), 187–197.

Riddel, M., 2001. A dynamic approach to estimating hedonic prices for environmental goods: an application to open space purchase. Land Economics 77 (4), 494–512.

Riley, J. G, Samuelson, W. F, 1981. Optimal auctions. American Economic Review 71, 381–392.

Rosen, S., 1974. Hedonic prices and implicit markets: product differentiation in pure competition. Journal of Political Economy 82 (1), 34–55.

Sander, H. A., Polasky, S., 2009. The value of views and open space: estimates from a hedonic pricing model for Ramsey County, Minnesota, USA. Land Use Policy 26, 837–845.

Schulz, T., Waltert, F., 2009. How local landscape resources affect property prices: evidence from a hedonic pricing model. Institute of Environmental Sciences, University of Zurich, working paper, Version 2009/03/01.

Shultz, S. D., King, D. A., 2001. The use of census data for hedonic price estimates of open-space amenities and land use. Journal of Real Estate Finance and Economics 22 (2/3), 239–252.

CHAPTER 3

A HEDONIC VALUATION FOR URBAN HOUSING WITH SPATIAL AND PROJECTS HETEROGENEITY: THE CASE OF CHENGDU, CHINA

Abstract

This study estimates the effect of spatial heterogeneity, project attributes and housing-unit attributes on the value of urban apartment housing using retail sales data in a Chinese housing market. To form the individual spatial weight matrix for each of the housing projects, we utilize the three-dimensional distances that not only take the plane coordinates into consideration, but also consider the floor on which the housing unit is located. With the aggregate spatial weight matrix transformed from the individual spatial weight matrices, we estimate both the spatial autoregressive model and the spatial error model using maximum likelihood. Our results show that for project attributes, the Plot Ratio and the weighted aggregate road service level have negative impacts on housing price, whereas subway station proximity and park proximity, as well as weighted aggregate healthcare service level, have positive impacts on housing price. In regard to housing-unit attributes, our results show that the coefficients of both inner and outer view variables are positive, while that of the "adjacent to a road" dummy variable is negative. In addition, our results confirm the positive impact of the direction of the major rooms in the housing unit when facing south, which is consistent with Chinese culture.

1. Introduction

This study estimates the effects of project attributes and housing-unit attributes on housing retail unit sales price. Our study uses data obtained from a Chinese regional housing market which consists of housing units in different housing projects. Unlike the "sparse" residential development pattern common in the US and other countries, the style of residential development in China is relatively more concentrated and denser. In fact, many large cities in Asia develop in a similar manner, and their residential buildings have the "high-rise" shape. Good examples are Hong Kong and Singapore. In the past 20 years, the residential development pattern in mainland China has become more and more dense as high-rise residential development has expanded from the coastal to the inland region, and is currently the prevalent urban development pattern. In contrast, due to its relatively large endowment of land, residential development in the US is much less dense, but some large cities such as downtown New York City and Chicago still have many high-rise residential buildings. The high-rise residential pattern presents a challenge to the traditional spatial hedonic approach since the standard two-dimensional concept in space does not fit the situation well. To our knowledge, no study has been done to conduct the hedonic estimation with respect to the high-rise residential pattern.

The typical pattern of residential development in China is that a real estate developer purchases a land parcel from the local government, and then builds several residential apartment buildings on the parcel.[20] There could be hundreds to thousands of housing units in a single housing project, depending on the size of the land parcel as well as the regulation requirement on its development density. In fact, given the large and dense housing projects in China, these projects often play a similar role as an entire community in the US. Large housing projects in China usually contain various kinds of open space amenities, sports fields, grocery stores, restaurants, and even kindergartens. Most of these projects are isolated by walls or fences, so that only residents and their invited guests can enter the housing projects. Housing projects in China, therefore, are analogous to "closed communities" in the US.

For each housing unit within a housing project, we primarily consider two types of explanatory variables that could affect its sales price: project attributes and housing-unit attributes.[21] For the first category of attributes, we consider those characteristics that could affect all the housing units within a particular housing project. Specifically, we examine the development density of the housing project, the proximity of the project to the nearest subway station and public park, and the overall healthcare service level as well as the service level of urban road network. For the second category of attributes, we consider those characteristics that could affect the

---

20 In China, all land belongs to the government. The maturity of the residential developable land is 70 years.

21 The housing sales price in China is usually listed in according to ￥/ m^2, not the total price per unit.

individual housing unit within each housing project, e.g., whether the housing unit has a view of an open-space amenity either within a housing project or outside the project, whether the housing unit is adjacent to a road or street, the unit's floor, the direction faced by the major rooms of the housing unit, the area of the housing unit, payment method, and the long term trend of the housing price.

The paper proceeds as follows: first we present a brief review of the traditional hedonic theory literature and its extension in a spatial context. After discussing the housing projects in Chengdu, we introduce the Relative Plane Coordinates System, and then we present our data. We then discuss some basic spatial hedonic models followed by discussion of the aggregate spatial weight matrix generated by the three-dimensional distances, which is a key feature of this study. Finally, we present our estimation results using both spatial autoregressive model and spatial error model which are estimated by the maximum likelihood approach.

2. Literature Review

Hedonic pricing studies date back to the pioneering works of Lancaster (1966), Ridker and Henning (1967), among others. Since the publication of Rosen's (1974) theoretical model, hedonic theory has been widely used in valuing the impact of environment and infrastructure on property values. The hedonic approach has been used to measure the changes of marginal willingness to pay in environmental attributes. Palmquist (1992) argues that marginal prices can measure total benefits sufficiently when externalities are localized. In addition, the hedonic approach uses data from real market transactions which can control for the hypothetical bias commonly found in the stated preference methods. The original hedonic approaches were used to value air quality; others looked at school quality, open space, mosquito abatement, road conditions, etc. In this research, we are primarily interested in how certain local public goods (i.e., open space and local infrastructure) and housing-unit attributes influence the housing price.

Open space is broadly defined as parks, rivers, or undeveloped land. In this study, our primary interest is public parks located within the main urban area of a city (see for example, Bolitzer and Netusil, 2000). While most studies consider the distance from a property to the source of open space, which is normally referred to as proximity, others have combined a measure of proximity with a measure of size, such as the area of a park (see Bolitzer and Netusil, 2000, and Mahan et al., 2000, for examples).[22] Other studies have also used a simple dummy variable to identify nearby open space amenities (see for example, Asabere and Huffman, 2009). In this study, we use a combination of these three approaches where applicable.

Gibbons and Machin (2003) and Nelson (1982) are good examples of studies examining the impact of a transportation system on property value. Gibbons and Machin (2003) use a method based on property values to evaluate the economic

22 While most distances are measured from centroid to centroid, there are a few studies that measure the distance from centroid to edge (see for Mahan et al., 2000, and Shultz and King, 2001).

benefits of transport access and transport innovation. They point out two benefits associated with the accessibility of rail: one is saving on travel times, and the other is the changes of the distribution of job types and wages. Essentially, easy access to a rail system can reduce the commuting costs to a great enough extent that potentially more-productive and higher-paid city jobs can be accessed. They define two ways to access a rail system: one is related to the distance to a station, and the other is the service frequency at the nearest station. They find proximity to a railway station and increased frequency positively affect property values. In addition, Nelson (1982) reviews nine studies on the effect of highway noise, finding that highway noise would cause a belt of roughly 1,000 feet that could negatively affect the nearby property value.

In addition to open space amenities and local infrastructure, property values are also regressed on various structural characteristics of the housing units, such as area of the unit, number of bedrooms, etc. (see for example, Lutzenhiser and Netusil, 2001 and Provencher et al., 2008). A variable capturing the "view" is commonly called "*View variable*." Sander and Polasky (2009) define the "*View variable*" in the following manner: viewshed area in square meters, standard deviation of elevations in a viewshed (measure of relief), view richness calculated as percentage of possible land use and land cover types contained in a viewshed, a viewshed composed of forest, a viewshed composed of grassy land covers, a viewshed composed of water, and a dummy variable indicating if a property has a view of downtown. Their results show that proximity to lakes has the greatest impact on home sale value. In addition, they find that view areal extents and the amount of water and grassy land covered in views also have positive impact on sale prices.

3. Market Setting and Data

3.1. The City of Chengdu and the Housing Projects

Our data set consists of six housing projects in Chengdu, China.[23] We have 1,268 observations (housing units) contained in six different housing projects (211 housing units per project on average). The city of Chengdu is the capital of Sichuan Province, which lies in the southwestern part of mainland China, about 1500 kilometers southwest of Beijing. It is situated at the western edge of the Sichuan Basin, with nearly 13 million official residents. Chengdu has the shape of a standard monocentric city. The most urbanized part of the city is surrounded by four concentric ring roads, with a fifth ring road under construction. Besides the ring roads, many radius roads also connect the center of the city to its edge in all directions. Currently, there are two subway lines being constructed from the north to the south, and from the west to the east, across the city. Our six housing projects are scattered across the city. Four housing projects are either within or around the third ring road, whereas the other two are further from the center of the city (see Fig.

23 We obtain sales data in the housing retail market from a local real estate sales agency, Chengdu SAGA Organization Ltd.

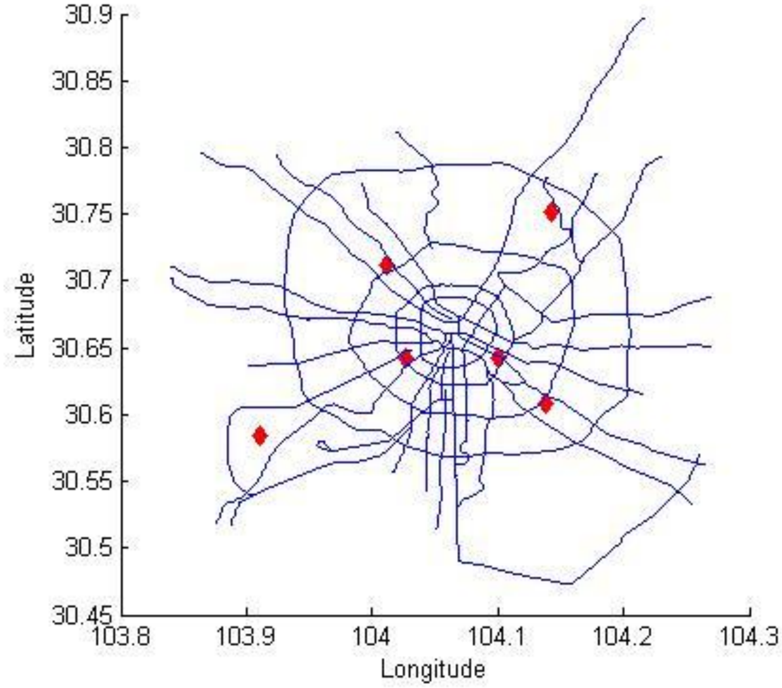

Fig. 3.1. Location of the Housing Projects

3.1 for the location of the housing projects).

3.2. Relative Plane Coordinates System

Unlike those commonly seen in the related literature, the information we have does not allow us to geo-code the housing units in each of the six housing projects using the GPS coordinates, because we do not have access to an up-to-date satellite image. In the absence of a GPS coordinate system (see for example, Anderson and West, 2006, and Bolitzer and Netusil, 2000), some researchers use a "grid" to geo-code the observations (see for Mahan et al., 2000). Since the housing units in the US commonly situate in a relatively sparse manner from one to another, the "grid" method of geo-coding works fairly well. In our case, however, housing units within one housing project are very dense, thus if we simply use a city-wide "grid" to geo-code these housing units we stand to lose a great deal of accuracy.[24]

In this study, we rely upon a "Relative Plane Coordinates System." There are two steps to implement this approach: first, we construct relative plane coordinates for each housing project; second, we calculate the length of the unit scale of each of the relative coordinates in meters. We have obtained the site plan of the housing projects from the local real estate sales agency along with the sales data. The sales agency has also assisted us with marking the room numbers of the housing units on the site plan. With this information we are able to use a simple but efficient way to geo-code the housing units.

Many graphic editing software programs have an auxiliary function called "ruler" which helps graphic designers locate elements in the graph more accurately. In our case, we use this ruler function to geo-code the housing units. The graphic software used is Photoshop.[25] One example is shown in Fig. 3.2. When we apply the ruler function, the software generates two rulers on both the top and left edges of the graph (the site plan in our case). These two rulers can play the role of a coordinates system.[26] After re-scaling the ruler distances, the distance between any two points

---

24 Some large housing projects may include 3,000 - 4,000 housing units, or more.
25 Researchers could use any other graphic software that has a "ruler" function.
26 Note that the origin generated by the software lies on the top-left corner,

Fig. 3.2. An Example of the Site Plan of the Housing Project with "Ruler"

in the plane is measured in meters.[27]

---

but we still have coordinates in a (x,y) pattern. To differentiate this system from one with height that we will discuss later, we add a term "plane" to it. This is why we call it "Relative Plane Coordinates System."

27 With this relative system that is not directly comparable for different housing projects, the scale changes due to the differentiated size of the graph. It is therefore necessary to transform each distance in different housing projects to a common scale (meters). We are able to accurately measure the distance of a given section along the edge in meters, $EDGE_i$, for $i=1,2,...,6$ denoting the 6 housing projects. Then, we turn to the site plan and find the corresponding two points along the edge of the site plan. Using the corresponding coordinates of these two points in our relative system, we calculate the distance between these two points under the relative plane coordinates system, denoted by $DIST_i$, for

3.3. Data Description

We divide the variables in our data set into three categories: Project-Attribute Variables, Housing-Unit-Attribute Variables, and Other Variables. For Project-Attribute Variables, we consider five variables: Plot Ratio, Subway Station Proximity, Park Proximity, Hospital Aggregate, and Road Aggregate. Plot Ratio is the ratio of total floor area to the land parcel area, which is the major index of development density. With respect to local infrastructure, we consider subway station proximity, healthcare service level (i.e., public hospitals), and major urban road network service level. Whether to use "proximity" or "level" is not only an empirical issue, but also a practical issue in the sense that we need to select those variables which are consistent with common sense.

For example, intuitively a measure of the distance to the nearest subway station would seem a more appropriate measure than a count of how many subway stations are surrounding a property. Subway Station Proximity, therefore, measures the Haversine distance from the centroid of each housing project to the nearest subway station.[28] Park Proximity measures the distance to the nearest public park in a

---

$i = 1,2,...,6$. We then calculate the unit scale of the relative plane coordinates system, $SCALE_i$, simply by: $SCALE_i = EDGE_i / DIST_i$, for $i = 1,2,...,6$. With this unit scale we can transform (standardize) each distance calculated in the relative plane coordinates system into meters by a simple multiplication, which is then comparable among different housing projects across the city.

28 The Haversine formula calculates the distance between any two points on a sphere. Haversine distance is usually obtained in the following steps: R = earth's radius (mean= 6,371km), $\Delta lat = lat_2 - lat_1$, $\Delta long = long_2 - long_1$, $a = \sin^2(\Delta lat/2) + \cos(lat_1) \times \cos(lat_2) \times \sin^2(\Delta long/2)$, $c = 2 \times \arcsin\{\min[1, sqrt(a)]\}$, $d = R \times c$. Note that all angles are measured in

similar manner. Hospital Aggregate measures the weighted aggregate healthcare service level evaluated at the centroid of each housing project. We use the number of beds in one hospital as its service level, and the reciprocal of the Haversine distance between the hospital to the target housing project as its "weight."[29] Thus, our aggregate measure of K public hospitals associated with a housing project located at latitude u and longitude v is,

$$a(u,v) = \sum_{j=1}^{K} a_k/z_k \tag{3.1}$$

where $a(u,v)$ measures the service level of public hospital in aggregate, a_k is the number of beds in the k^{th} hospital, and z_k is the distance from the housing project to the k^{th} hospital. In addition, Road Aggregate is calculated in a similar manner. We depict the major road network in the city by 70,826 points with the GPS coordinates (see Fig. 3.1). We assign a unit "1" to all the road location points as their "level," and follow exactly the same approach as that for the hospitals to calculate the weighted aggregate road service level. The descriptive statistics of all the five variables in the category of Project attributes are reported in Table 3.1.

Sander and Polasky (2009) calculate viewsheds using the *VIEWSHED* function in a software called ArcGIS. Unfortunately, we do not have access to local GIS database that could satisfy the requirement of the *VIEWSHED* function in ArcGIS. We therefore only use the dummy variable approach to represent the "*View variables*"

---

radians.

29 We consider all the public hospitals in the city.

Table 3.1
Descriptive Statistics of Variables in Project Attributes

	Mean	Median	Std. Dev.
Plot Ratio	4.354	4.380	2.723
Subway Station Proximity (m)	5903.838	3215.304	5535.142
Park Proximity (m)	1348.407	1305.554	484.032
Hospital Aggregate	0.082	0.065	0.040
Road Aggregate	8.132	7.466	2.695

in this study. In addition, due to the distinct feature of the Chinese culture, we also introduce the "*Direction variables*," which could be considered as special variants of the standard "*View Variables*." In regard to the Housing-Unit-Attribute Variables in this study, the first thing we need to consider is the "*View*" of the housing units to either the major open-space amenity source within each housing project or open-space amenity source that is outside but adjacent to the housing project.

As the site plan in Fig. 3.2 shows, the major open-space amenity sources within this small housing project are the swimming pool and some small gardens nearby, which lie in the center of the project surrounded by the residential buildings. We assign a value "1" to those housing units that are able to see the swimming pool. In addition, to the north of this housing project, a small public park (the green land as shown in the site plan) is an open-space amenity source outside but adjacent to this housing project. Again, for those housing units that have a view of this public park, we assign a value "1."

In addition to these two "amenity" view variables, we also introduce a "disamenity" dummy variable to show whether a housing unit is close to the urban road / street. Living close to a major urban road (especially directly facing it), residents would suffer from noise and dust. We therefore expect a negative sign for the coefficient on this dummy variable. Note that these three "view" variables are not mutually exclusive.[30] Thus, we include each of these three variables in our estimation. In our data set, 39.2% of the housing units have view to major inner open space amenities, 23.1% of the housing units have view to open space amenities right outside the housing projects, and 31.9% of the housing units are located on the fringe of the housing projects which are close to urban roads and streets.

In the Chinese culture, people pay attention to the direction of the major rooms (such as living room, main bedroom, etc.) when they choose the location of their housing units. Facing south is considered to be the most preferable direction for living, which is believed to make the room cool in summer and warm in winter. We therefore use eight dummy variables to depict the direction of the housing units. In our data set, 14% of the housing units have their major rooms facing directly to the South (see Table 3.2). Since the dummy variables of these eight directions are mutually exclusive, we omit "Northwest," which has the most observations in the data set.

Other variables in the category of housing unit attributes include "Floor," "Unit

---

30 In fact, 14.8% of the observations in our data set have two of these three attributes at the same time; 6.3% of the observations have all the three attributes.

Table 3.2
Descriptive Statistics of Variables in Housing-Unit Attributes and Dependent Variables

	Mean	Median	Std. Dev.
Floor	10.625	9.000	7.540
Unit area (m^2)	95.125	89.370	31.965
Distance to major open-space amenity within each project (m)	176.843	97.761	229.999
Time trend (1= Jan, 2004)	61.217	62.000	6.9703
Inner view (1=yes)	39.236%	--	0.493
Outer view (1=yes)	23.090%	--	0.445
Close to street (1=yes)	31.858%	--	0.478
North (1=yes)	13%	--	0.319
North East (1=yes)	7%	--	0.246
East (1=yes)	11%	--	0.300
South East (1=yes)	17%	--	0.386
South (1=yes)	14%	--	0.337
South West (1=yes)	13%	--	0.330
West (1=yes)	5%	--	0.204
North West (1=yes)	20%	--	0.427
Deflated unit sales price (￥/ m^2)	4130.603	4256.436	1256.344
Non-deflated unit sales price (￥/ m^2)	5551.158	5685.0858	1658.371

area," and "Distance to major open-space amenity within each project." "Floor," which shows the number of stories at which the housing unit situates, in fact gives the height of the housing unit. "Unit area" is not only a quantitative index, but also a qualitative index. Normally in the market, the larger the area of the housing unit, the more luxurious it is considered to be. Thus, by treating "Unit area" as a characteristic of the housing unit, we expect a positive sign for its coefficient. "Distance to major open-space amenity within each project" is a key explanatory variable in this study. Using our relative plane coordinates system, we obtain the coordinates for each of the housing units, along with the coordinates of the centroid of the open-space amenity source within each housing project. With these coordinates, we could calculate the Euclidian distance from each housing unit to the source of the in-project open space. Note that these distances are two dimensional, so that the distance from a housing unit at the top of a building to the open space source would be the same as that at the bottom of the building.

Besides the variables in these two categories, we consider two additional other variables: payment method and time trend. The first shows the choice of payment method. Normally in China, when consumers purchase housing with cash, not a mortgage, they receive some discount from the developer. Therefore, we would expect a negative sign for the coefficient of this variable. In our data set, 73.5% of the housing units are purchased via mortgage.

Sales for the entire housing project usually takes a long period of time,[31] and the 6 housing projects in our data set were not marketed during the same time period, thus we need to consider the issue of time in our study. Common approaches to deal with the time issue are: deflation by a given price index,[32] such as housing price

---

31 In China, the developers usually sell housing units in "batches," thus it may take several years for large housing projects to complete their development and sales.
32 See for example, Bolitzer and Netusil (2000) and Lutzenhiser and Netusil (2001).

index (HPI), as well as the use of certain time dummy variables.[33] In this study, we also consider a use of a mix of deflation by HPI and a time dummy. The choice of these three approaches is just an empirical issue, therefore, we use all of them in our study. We set January 2004 to be 1, and use a step size of 1 for every additional month.[34]

Against each explanatory variable that discussed above, we set the variable of "Deflated unit sales price" as our dependant variable.[35] We use a local monthly housing price index (HPI) to deflate the short term fluctuation in the actual housing sales price.[36] We set the value of the HPI to be 1 in January 2004, and then divide the actual sales price by the corresponding value of HPI. Descriptive statistics of all the variables in the second and the third categories along with the dependent variables are shown in Table 3.2.

4. Empirical Models

4.1. Spatial Hedonic Models

Rosen's (1974) seminal article proposed an equilibrium model of product

33 See for example, Provencher et al. (2008). The annual dummy variables are included to represent the temporal shifts in the residential property market.
34 Note that, for the "deflation only" case, we do not include the monthly dummy variables.
35 Note that, for the case of "dummy only," we do not use the deflated unit sales prices as the dependent variable. Instead, we directly use the non-deflated sales price.
36 The housing price index in Chengdu is reported monthly by the National Development and Reform Commission.

differentiation. In a competitive market, goods are valued for their utility-bearing characteristics, and the interactions of buyers and sellers over multiple attributes yield the hedonic equilibrium price function. On the supply side, a set of sellers propose several price schemes over the vector of characteristics *z*, known as the "offer curve," $\phi(z)$. On the demand side, many buyers also propose a set of valuation on *z*, known as the "bid curve," $\theta(z)$. The price function $P(z)$ is hence an envelope of the tangent points between the offer and bid functions, as Fig. 3.3 shows.

Since the housing retail market in our study can be considered competitive, traditional hedonic theory applies. However, a commonly seen econometric problem in hedonic modeling is that the data may be spatially correlated. Anselin (1988) has shown that OLS estimates are inconsistent under spatial heterogeneity. A quick examination of our data confirms strong spatial correlation.[37] Although there are several variants, Anselin (1988) has discussed the two fundamental spatial econometric models that address the spatial correlation problem — the spatial-lag (SAR) and spatial-error (SEM) models. The functional form of the SAR is,

$$y = \boldsymbol{\rho} \times W \times y + X \times \boldsymbol{\beta} + e \tag{3.2}$$

whereas the form of the SEM is given by,

$$y = X \times \boldsymbol{\beta} + u,\ u = \boldsymbol{\lambda} \times W \times u + e \tag{3.3}$$

37 We used test statistics for Moran's I-test, a likelihood ratio test, a Wald test, and a Lagrange multiplier test for spatial correlation in the residuals, and a Lagrange multiplier test for correlation in the SAR residuals. See Anselin (1988) and LeSage (1999) for details. All five tests show strong spatial correlation in the housing unit sales price.

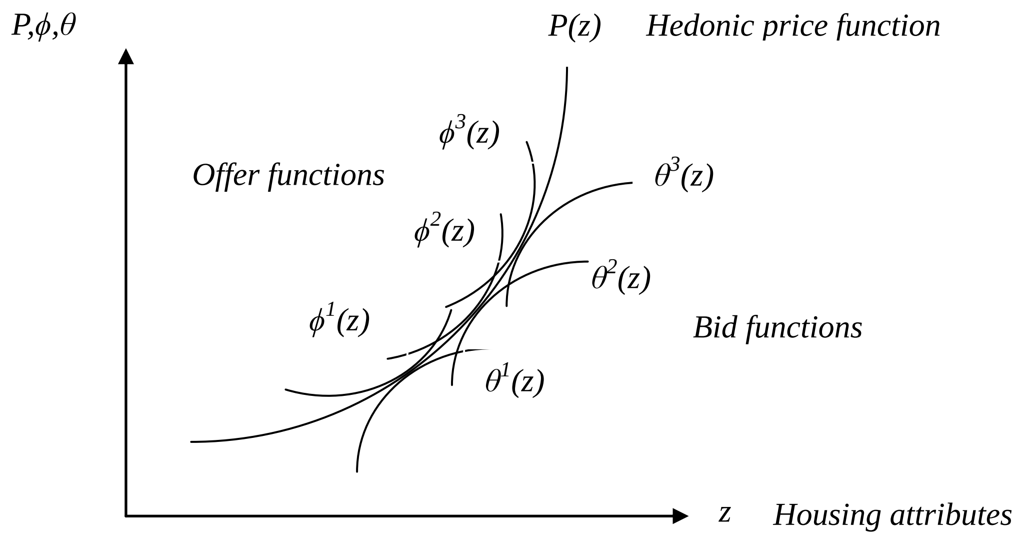

Fig. 3.3. Hedonic Price Function in the Housing Retail Market

The most important features of these models in Eqs. (3.2) and (3.3), are W, $\boldsymbol{\rho}$, and $\boldsymbol{\lambda}$, where W is known as the spatial weight matrix, $\boldsymbol{\rho}$ and $\boldsymbol{\lambda}$ are the spatial lag coefficients in both SAR and SEM, respectively. The maximum likelihood estimation (MLE) approach is the most popular estimation method for the spatial models. As Kim et al. (2003) have shown, spatial techniques in hedonic pricing models usually increase both the R^2 and the significance of estimated coefficients. Further discussion of spatial hedonic models can be found in Irwin (2002) and others.

4.2. Spatial Weight Matrix

There are two difficulties in forming the spatial weight matrix in our study. First, as normally seen in literature, the spatial weight matrix is constructed in a plane, two-dimensional surface. However, a 2-D spatial weight matrix does not apply to our data.[38] Since the high-rise residential buildings in our data set have multistories, it is possible that the sales price of two housing units with similar characteristics but different height are spatially correlated. However, in a 2-D setting, the distance between these two units would be zero, which fails to capture the spatial nature in height. Therefore, we need to extend the traditional 2-D spatial weight matrix into a three-dimensional (3-D) setting.

Before we formally construct our spatial weight matrix, we need to introduce the three-dimensional distance. As Fig. 3.4 shows, the 3-D distance in a (x,y,z) coordinates system is given by line *BC*, rather than line *OB* as the 2-D distance in a (x,y) coordinates system. The 3-D distance *BC*, therefore, is given as follows:[39]

$$BC = [(OA^2 + AB^2)^{1/2} + OC^2]^{1/2} \tag{3.4}$$

Once we have obtained the 3-D distances, we follow the standard approach to form the spatial weight matrix for each of the six housing projects, i.e., use the reciprocal of the 3-D distances as the elements of the spatial weight matrix. We denote these six spatial weight matrices by $W_1, W_2, \ldots, W_6$.

The second difficulty is how to form an "aggregate" spatial weight matrix. Since our relative coordinates system described previously is not a global coordinates

---

38 In fact, in our preliminary estimation, the 2-D spatial weight matrix has done a very poor job. The matrix is nearly singular and cannot be inverted.
39 In the local market, the height of each floor is commonly 3 meters, we therefore calculate the height of the housing unit, *OC*, approximately by: $OC = 3 \times (Floor - 1)$.

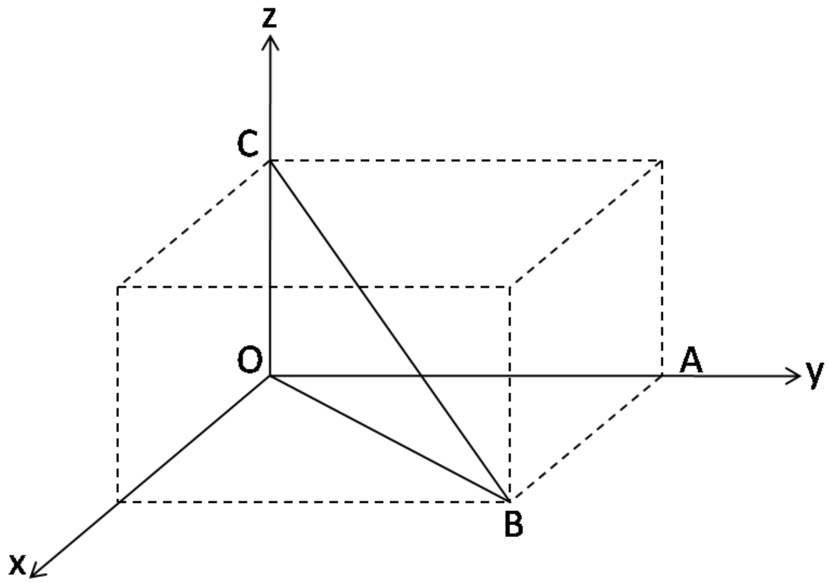

Fig. 3.4. Three-Dimensional Distance

system such as GPS, a project-specific spatial weight is not directly comparable to that of another project. Each weight matrix is valid only within the limit of the individual housing project. However, we could use two approaches to depict the relationship of the housing units among different housing projects. Perhaps the easiest and most straight forward way is to consider the impact of the housing units in one housing project on housing units in other housing projects to be zero. While the distances between different housing units within a housing project are small (usually measured in meters), the distances between different housing projects are very large (usually measured in kilometers). We could therefore let the reciprocal of the distances of the housing units at different housing projects equal to zero. Paying attention to the dimensions of the elements, we could construct the "aggregate" spatial weight matrix as follows:

$$W = \begin{bmatrix} W_1 & O_{(n1,n2)} & O_{(n1,n3)} & O_{(n1,n4)} & O_{(n1,n5)} & O_{(n1,n6)} \\ O_{(n2,n1)} & W_2 & O_{(n2,n3)} & O_{(n2,n4)} & O_{(n2,n5)} & O_{(n2,n6)} \\ O_{(n3,n1)} & O_{(n3,n2)} & W_3 & O_{(n3,n4)} & O_{(n3,n5)} & O_{(n3,n6)} \\ O_{(n4,n1)} & O_{(n4,n2)} & O_{(n4,n3)} & W_4 & O_{(n4,n5)} & O_{(n4,n6)} \\ O_{(n5,n1)} & O_{(n5,n2)} & O_{(n5,n3)} & O_{(n5,n4)} & W_5 & O_{(n5,n6)} \\ O_{(n6,n1)} & O_{(n6,n2)} & O_{(n6,n3)} & O_{(n6,n4)} & O_{(n6,n5)} & W_6 \end{bmatrix} \tag{3.5}$$

where O is zero matrix, and $n_1,n_2,...,n_6$ stand for the number of observations in each housing project. An alternative method of forming the aggregate spatial weight matrix is to use the reciprocal of the distances among different housing projects (as measured from centroid to centroid) as the off-diagonal elements. Specifically, we would construct the "aggregate" spatial weight matrix in the following manner:

$$W = \begin{bmatrix} W_1 & \frac{1}{d_{12}}\times I_{(n1,n2)} & \frac{1}{d_{13}}\times I_{(n1,n3)} & \frac{1}{d_{14}}\times I_{(n1,n4)} & \frac{1}{d_{15}}\times I_{(n1,n5)} & \frac{1}{d_{16}}\times I_{(n1,n6)} \\ \frac{1}{d_{21}}\times I_{(n2,n1)} & W_2 & \frac{1}{d_{23}}\times I_{(n2,n3)} & \frac{1}{d_{24}}\times I_{(n2,n4)} & \frac{1}{d_{25}}\times I_{(n2,n5)} & \frac{1}{d_{26}}\times I_{(n2,n6)} \\ \frac{1}{d_{31}}\times I_{(n3,n1)} & \frac{1}{d_{32}}\times I_{(n3,n2)} & W_3 & \frac{1}{d_{34}}\times I_{(n3,n4)} & \frac{1}{d_{35}}\times I_{(n3,n5)} & \frac{1}{d_{36}}\times I_{(n3,n6)} \\ \frac{1}{d_{41}}\times I_{(n4,n1)} & \frac{1}{d_{42}}\times I_{(n4,n2)} & \frac{1}{d_{43}}\times I_{(n4,n3)} & W_4 & \frac{1}{d_{45}}\times I_{(n4,n5)} & \frac{1}{d_{46}}\times I_{(n4,n6)} \\ \frac{1}{d_{51}}\times I_{(n5,n1)} & \frac{1}{d_{52}}\times I_{(n5,n2)} & \frac{1}{d_{53}}\times I_{(n5,n3)} & \frac{1}{d_{54}}\times I_{(n5,n4)} & W_5 & \frac{1}{d_{56}}\times I_{(n5,n6)} \\ \frac{1}{d_{61}}\times I_{(n6,n1)} & \frac{1}{d_{62}}\times I_{(n6,n2)} & \frac{1}{d_{63}}\times I_{(n6,n3)} & \frac{1}{d_{64}}\times I_{(n6,n4)} & \frac{1}{d_{65}}\times I_{(n6,n5)} & W_6 \end{bmatrix} \tag{3.6}$$

where d_{ij} is the distance from housing project i to housing project j, for $i,j=1,2,...,6$; $I_{(ni,nj)}$ is the identity matrix.

Once we have decided upon a spatial weight matrix, we can implement spatial econometric techniques. We first use the aggregate spatial weight matrix defined in Eq. (3.5), and then try the one defined in Eq. (3.6). For both matrices, we estimate both the SAR and the SEM. The estimation results for the SAR are reported in Tables 3.3, 3.4 and 3.5, for the cases using HPI deflation, monthly time dummy, and both deflation and dummy, respectively. In addition, the results of SEM are reported in Tables 3.6, 3.7 and 3.8. The results using the aggregate spatial matrix defined in Eq. (3.6) are similar to those using the one defined in Eq. (3.5), thus we only present the results using the simpler aggregate spatial weight matrix defined in Eq. (3.5) here.[40]

4.3. Estimation Results

Overall, SEM appears to perform better than SAR, in the sense that several key explanatory variables are more statistically significant even though their signs are roughly the same. The adjusted R^2 of the SEM models are larger than the SAR models as well. Among the three SEM models, the one with only a monthly time dummy (Table 3.7) appears to have the largest R^2, however, the sign of "Park proximity" is not as expected, since a positive sign means that the further away from the park the higher of the housing price, which is inconsistent with our common sense. For the cases using HPI deflation (Table 3.6) and the one with both deflation and

40 Estimation results using the aggregate spatial weight matrix defined in Eq. (3.6) are available upon request.

Table 3.3
Spatial Autoregressive Model of Deflated Unit Sales Price, P_H (￥/ m^2), Using HPI Deflation

Variable	Coefficient	t-statistic	p-value
Intercept	1.1505	1.2566	0.2089
Project-Attribute Variables			
Plot Ratio	-0.0539	-1.2772	0.2015
ln(Subway station proximity)	-0.0825	-1.2554	0.2093
ln(Park proximity)	-0.0010	-0.1242	0.9012
Hospital aggregate	5.2838	1.4339	0.1516
Road aggregate	-0.0678	-1.4123	0.1578
Housing-Unit-Attribute Variables			
Floor	0.0001	0.4861	0.6269
Inner view (1=yes)	0.0370	9.4551	0.0000
Outer view (1=yes)	0.0232	4.7904	0.0003
Close to street (1=yes)	-0.0050	-1.1702	0.2419
North (1=yes)	-0.0119	-1.9070	0.0565
North East (1=yes)	-0.0037	-0.5106	0.6096
East (1=yes)	-0.0185	-2.7652	0.0057
South East (1=yes)	-0.0043	-0.9329	0.3509
South (1=yes)	0.0494	7.9155	0.0000
South West (1=yes)	0.0007	0.1132	0.9098
West (1=yes)	-0.0005	-0.0595	0.9526
Distance to inner source	-0.0001	-1.4480	0.1476
Housing unit area	0.0004	6.3184	0.0000
Other Variables			
Pay in cash (1=yes)	-0.0140	-4.2672	0.0000
$\boldsymbol{\rho}$	0.9790	126.2010	0.0000
Adjusted R-square		0.9335	
sigma^2		0.0023	
log-likelihood		2466.5871	

Dependent variable: ln(Deflated housing unit price)

Table 3.4
Spatial Autoregressive Model of Unit Sales Price, P_H (￥/ m^2), Using Monthly Time Dummy

Variable	Coefficient	t-statistic	p-value
Intercept	1.1929	1.2704	0.2039
Project-Attribute Variables			
Plot Ratio	-0.0553	-1.2772	0.2015
ln(Subway station proximity)	-0.0850	-1.2622	0.2069
ln(Park proximity)	-0.0014	-0.1232	0.9020
Hospital aggregate	5.4235	1.4437	0.1488
Road aggregate	-0.0670	-1.4320	0.1521
Housing-Unit-Attribute Variables			
Floor	0.0002	0.8121	0.4167
Inner view (1=yes)	0.0372	9.3948	0.0000
Outer view (1=yes)	0.0238	4.8756	0.0000
Close to street (1=yes)	-0.0049	-1.1374	0.2554
North (1=yes)	-0.0122	-1.9305	0.0536
North East (1=yes)	-0.0054	-0.7324	0.4639
East (1=yes)	-0.0170	-2.9451	0.0032
South East (1=yes)	-0.0059	-1.2650	0.2059
South (1=yes)	0.0484	7.6454	0.0000
South West (1=yes)	0.0000	0.0065	0.9948
West (1=yes)	-0.0006	-0.0674	0.9463
Distance to inner source	-0.0001	-1.4880	0.1368
Housing unit area	0.0004	6.1894	0.0000
Other Variables			
Pay in cash (1=yes)	-0.0135	-4.0578	0.0001
Time trend	-0.0000	-0.0243	0.9806
$\boldsymbol{\rho}$	0.9790	124.7119	0.0000
Adjusted R-square		0.9321	
sigma^2		0.0024	
log-likelihood		2453.2726	

Dependent variable: ln(Housing unit price)

Table 3.5
Spatial Autoregressive Model of Deflated Unit Sales Price, P_H (￥/ m^2), Using Both HPI Deflation and Monthly Dummy

Variable	Coefficient	t-statistic	p-value
Intercept	1.2697	1.4208	0.1554
Project-Attribute Variables			
Plot Ratio	-0.0551	-1.3575	0.1746
ln(Subway station proximity)	-0.0858	-1.3569	0.1748
ln(Park proximity)	-0.0008	-0.0671	0.9465
Hospital aggregate	5.3425	1.5204	0.1284
Road aggregate	-0.0679	-1.4918	0.1358
Housing-Unit-Attribute Variables			
Floor	0.0001	0.5004	0.6168
Inner view (1=yes)	0.0371	9.4538	0.0000
Outer view (1=yes)	0.0233	4.8107	0.0000
Close to street (1=yes)	-0.0051	-1.2011	0.2297
North (1=yes)	-0.0119	-1.9048	0.0568
North East (1=yes)	-0.0037	-0.5023	0.6155
East (1=yes)	-0.0186	-2.7728	0.0056
South East (1=yes)	-0.0043	-0.9227	0.3561
South (1=yes)	0.0496	7.9120	0.0000
South West (1=yes)	0.0008	0.1298	0.8967
West (1=yes)	-0.0004	-0.0436	0.9652
Distance to inner source	-0.0001	-1.5345	0.1249
Housing unit area	0.0005	6.4325	0.0000
Other Variables			
Pay in cash (1=yes)	-0.0140	-4.2615	0.0000
Time trend	-0.0000	-0.0806	0.9358
$\boldsymbol{\rho}$	0.9680	104.2122	0.0000
Adjusted R-square		0.9509	
sigma^2		0.0023	
log-likelihood		2463.8601	

Dependent variable: ln(Deflated housing unit price)

Table 3.6
Spatial Error Model of Deflated Unit Sales Price, P_H (￥/ m^2), Using HPI Deflation

Variable	Coefficient	t-statistic	p-value
Intercept	14.4969	52.9346	0.0000
Project-Attribute Variables			
Plot Ratio	-0.3005	-4.7624	0.0001
ln(Subway station proximity)	-0.5952	-6.5643	0.0000
ln(Park proximity)	-0.0116	-0.1106	0.9119
Hospital aggregate	21.4001	3.5758	0.0003
Road aggregate	-0.2268	-3.3759	0.0007
Housing-Unit-Attribute Variables			
Floor	-0.0002	-0.5301	0.5961
Inner view (1=yes)	0.0433	9.0134	0.0000
Outer view (1=yes)	0.0229	3.5878	0.0003
Close to street (1=yes)	-0.0132	-2.4402	0.0147
North (1=yes)	-0.0165	-2.0895	0.0367
North East (1=yes)	-0.0145	-1.4672	0.1423
East (1=yes)	-0.0188	-2.2634	0.0236
South East (1=yes)	-0.0065	-1.1253	0.2604
South (1=yes)	0.0561	6.8528	0.0000
South West (1=yes)	0.0001	0.0156	0.9876
West (1=yes)	0.0053	0.5160	0.6058
Distance to inner source	-0.0003	-2.6386	0.0083
Housing unit area	0.0009	7.8160	0.0000
Other Variables			
Pay in cash (1=yes)	-0.0144	-4.1225	0.0000
λ	0.9660	899.8973	0.0000
Adjusted R-square		0.9736	
sigma^2		0.0023	
log-likelihood		2464.2008	

Dependent variable: ln(Deflated housing unit price)

Table 3.7

Spatial Error Model of Unit Sales Price, P_H (¥/ m^2), Using Monthly Time Dummy

Variable	Coefficient	t-statistic	p-value
Intercept	14.5140	42.5045	0.0000
Project-Attribute Variables			
Plot Ratio	-0.2808	-3.5207	0.0004
ln(Subway station proximity)	-0.5922	-5.1690	0.0000
ln(Park proximity)	0.0143	0.1076	0.9144
Hospital aggregate	19.2306	2.5450	0.0109
Road aggregate	-0.2085	-2.4660	0.0137
Housing-Unit-Attribute Variables			
Floor	-0.0001	-0.2433	0.8078
Inner view (1=yes)	0.0433	8.9128	0.0000
Outer view (1=yes)	0.0237	3.6710	0.0002
Close to street (1=yes)	-0.0132	-2.4085	0.0160
North (1=yes)	-0.0170	-2.1199	0.0340
North East (1=yes)	-0.0176	-1.7576	0.0788
East (1=yes)	-0.0205	-2.4467	0.0144
South East (1=yes)	-0.0086	-1.4736	0.1406
South (1=yes)	0.0544	6.5593	0.0000
South West (1=yes)	-0.0014	-0.1687	0.8661
West (1=yes)	0.0050	0.4847	0.6277
Distance to inner source	-0.0003	-2.5637	0.0104
Housing unit area	0.0009	7.6168	0.0000
Other Variables			
Pay in cash (1=yes)	-0.0138	-3.9076	0.0001
Time trend	0.0002	0.2483	0.8039
λ	0.9730	1012.4403	0.0000
Adjusted R-square		0.9748	
sigma^2		0.0024	
log-likelihood		2452.8514	

Dependent variable: ln(Housing unit price)

Table 3.8
Spatial Error Model of Deflated Unit Sales Price, P_H (￥/ m^2), Using Both HPI Deflation and Monthly Dummy

Variable	Coefficient	t-statistic	p-value
Intercept	14.4433	43.3222	0.0000
Project-Attribute Variables			
Plot Ratio	-0.2978	-3.8959	0.0001
ln(Subway station proximity)	-0.5934	-5.3998	0.0000
ln(Park proximity)	-0.0070	-0.0550	0.9562
Hospital aggregate	21.0537	2.9093	0.0036
Road aggregate	-0.2245	-2.7688	0.0056
Housing-Unit-Attribute Variables			
Floor	-0.0002	-0.5404	0.5889
Inner view (1=yes)	0.0432	9.0004	0.0000
Outer view (1=yes)	0.0228	3.5683	0.0004
Close to street (1=yes)	-0.0131	-2.4203	0.0155
North (1=yes)	-0.0164	-2.0726	0.0382
North East (1=yes)	-0.0147	-1.4784	0.1393
East (1=yes)	-0.0187	-2.2542	0.0242
South East (1=yes)	-0.0065	-1.1295	0.2587
South (1=yes)	0.0560	6.8248	0.0000
South West (1=yes)	0.0001	0.0070	0.9944
West (1=yes)	0.0052	0.5085	0.6111
Distance to inner source	-0.0003	-2.5423	0.0110
Housing unit area	0.0009	7.7069	0.0000
Other Variables			
Pay in cash (1=yes)	-0.0144	-4.1220	0.0000
Time trend	0.0001	0.1681	0.8665
λ	0.9720	995.0180	0.0000
Adjusted R-square		0.9737	
sigma^2		0.0023	
log-likelihood		2465.9607	

Dependent variable: ln(Deflated housing unit price)

dummy (Table 3.8), the results appear to be roughly the same. Since the one using both deflation and dummy has slightly larger R^2 and log likelihood values, we consider the SEM using a mix of deflation and dummy to be the preferable estimation model in this study.

As Table 3.8 shows, our model seems to work fairly well, in the sense that most of the key variables have "correct" sign and are statistically significant. The adjusted R^2 is very large as well. Among project attributes, Plot Ratio has a negative impact on the housing price, which implies that consumers do not like living in a very dense housing project, and developers must therefore accept a lower per unit price the greater the density. Subway station proximity has a strong positive impact on the housing price which indicates that consumers prefer to living close to a subway station, and developers can charge a premium on it. Park proximity is also found positively related to the housing price, however, it is not statistically significant. The healthcare service level (i.e., Hospital aggregate) has a significant positive influence on the housing price and its coefficient is very large (which shows that an increase of one unit of healthcare service level could increase about 21% of the housing unit price given our semi-log functional form, other things equal). But recall that the mean value of the healthcare service level is only 0.082 given our calculation method, thus it is not easy to increase the weighted aggregate hospital service level by one unit.[41] In addition, the weighted aggregate of urban road service level influences the housing price significantly in a negative manner, which shows that consumers do not like living at a location that is surrounded by a very dense road network.

In regard to housing unit attributes, the floor on which the housing unit is located has a statistically insignificant impact on the housing price. This is somewhat consistent with our expectations since, based on our knowledge of the local market, the highest housing unit sales price for a high-rise residential building usually occurs in the middle of the building, particularly when the housing unit faces the inside of the project; however, it increases as the housing unit is located on a higher floor and faces the outside of the project. Both the dummy variables of inner view and outer view show significantly positive influences on housing price. As expected, the dummy variable for adjacent to a urban road shows a significantly negative impact on housing price.

In our estimation results, compared to the omitted directional dummy variable for facing Northwest, we have identified a strong positive impact of facing South on the housing price, which is consistent with our expectations. The other six directions have

---

41 Recall that the healthcare service level is calculated as the summation of number of beds in hospitals weighted by the distances. Given the distances (while a household can choose to locate close to a particular hospital, it is impossible to live close to all the hospitals), it would be difficult to increase the number of beds in the hospitals to a large degree. Hence, if the healthcare service level increases from 0.082 to 0.092 (a 0.01 level change), the effect on the per unit housing price is only 21% × 0.01 = 0.21%.

either positive or negative influences on the housing price. In regard to the plane distance to the major open-space amenity within each housing project, we have found that the closer the housing unit lies to a source of major "in-project" open-space amenity, the higher is the housing price.

For the housing unit area, though, the estimated coefficient is small, its *t* statistic is large, which implies price increases as the living area increases. We know that the total price of a housing unit equals its unit sales price times area. Our finding indicates that the size of the housing unit positively and significantly affects unit sales price. What exactly is the premium?

Consider two housing units of sizes 95 m^2 and 110 m^2, respectively. Assuming that the unit sales price of the 95 m^2 one is 5551￥/m^2.[42] According to our estimation results, a 0.0135% increase due to the 15 m^2 change in area would result in an increase of 74.9￥/m^2 in the unit sales price, to 5626￥/m^2. For the 110 m^2 housing unit, the total premium due to a 15 m^2 increase in size would be ￥84,390.[43]

For the remaining explanatory variables, we find that *pay-in-cash* significantly lowers housing price compared to the mortgage payment method, which is consistent

42 In this example, both the area of the housing unit and its unit sales price are the mean values in our data set.

43 Denoting the total price of a housing unit as *TP*, since the unit sales price P_H is a function of the housing area M, we have: $TP = P_H(M) \times M$. Totally differentiate *TP* with respect to *M*, we have:
$\left.\frac{\partial TP}{\partial M}\right|_{M0} \times dM = [\left.\frac{\partial TP}{\partial M}\right|_{M0} \times M_0 + P_H(M_0)] \times dM$. In our example, $M_0 = 95$, $P_H(M_0) = 5551$, and $dM = 15$.

with our expectations. For the long term time trend, our estimation results show that there is a slightly positive but statistically insignificant trend of the housing price in the long run, which reveals that the short term fluctuation that has been taken out by deflating it with the monthly housing price index (HPI) explains the time effect of the local housing market relatively well. In addition, the spatial lag coefficient of the error term is large and statistically significant. This comes as no surprise since housing units within a given housing project of the high-rise pattern are very close to each other, which is the typical situation for residential buildings in China.[44]

5. Concluding Remarks

In this study, we have estimated the effect of spatial heterogeneity in project and housing-unit attributes on the value of urban housing using retail sales data in a Chinese regional housing market. We use three-dimensional distances that consider the floor on which the housing unit is located to form the individual spatial weight matrix for each housing project.

Our estimation results show that for project attributes, the Plot Ratio and the weighted aggregate road service level have negative impacts on housing price, whereas subway station proximity and park proximity, as well as weighted aggregate healthcare service level, have positive impacts on housing price. In regard to

44 Imagine that your housing unit is surrounded by your neighbors from front, back, left, and right, as well as top and bottom.

housing-unit attributes, our results show that the coefficients of all the view variables (both inner and outer) are positive, while the coefficient of the dummy variable showing whether the housing unit situates adjacent to a road or street is negative. In addition, our results confirm the positive impact of the direction of the major rooms in the housing unit when facing south, which is consistent with Chinese culture. We have found evidence of a positive impact of the housing unit area on its unit sales price, and we have obtained an insignificantly negative coefficient of the floor variable, which is also consistent with our expectation of the complex relationship between the floor and the unit sales price. For other variables, although we have found that payment method affects the unit sales price, our results do not suggest that there is a significant long term time trend in this price.

This study contributes to the literature not only for its unique data set, but also for aggregate weight matrix based on 3-D distance. However, this study could be extended in two ways. One is the identification issue. As Palmquist (1984) has pointed out, the cost of moving between cities would block the complete integration of the property markets between cities. Thus, the hedonic price parameters may vary in different markets even for the same set of explanatory variables. Therefore, if we could obtain housing retail sales data in other cities in China (especially in those large cities in the southwestern region, such as Chongqing, Kunming, Guizhou, etc.), then we are able to compare the estimation results in this study with those at other markets.

Another issue is market segmentation. Although we consider the local housing market to be competitive, in fact the property developer has some pricing power within the boundary of its housing project. Therefore, there may be a single offer function in each housing project. Hence the estimation parameters may vary among different housing projects. Therefore, if we could test whether the estimation parameters are equal across different housing projects, we would be able to identify possible market segmentation in the same regional housing market.

References

Anderson, S. T., West, S. E., 2006. Open space, residential property values, and spatial context. Regional Science and Urban Economics 36, 773–789.

Anselin, L., 1988. Spatial econometrics: methods and models. Kluwer Academic Press. Dordrecht.

Asabere, P. K., Huffman, F. E., 2009. The relative impacts of trails and greenbelts on home price. Journal of Real Estate Finance and Economics 38, 408–419, doi: 10.1007/s11146-007-9089-8.

Bolitzer, B., Netusil, N. R., 2000. The impact of open spaces on property values in Portland, Oregon. Journal of Environmental Management 59, 185–193, doi: 10.1006/jema.2000.0351.

Gibbons, S., Machin, S., 2003. Rail access and house prices: an evaluation of the wider benefits of transport. Department of Transport Report , United Kingdom (November).

Irwin, E. G., 2002. The effects of open space on residential property values. Land Economics 78 (4), 465–480.

Kim, C. W., Phipps, T. T., Anselin, L., 2003. Measuring the benefits of air quality improvement: a spatial hedonic approach. Journal of Environmental Economics and Management 45, 24–39.

Lancaster, K. J., 1966. A new approach to consumer theory. Journal of Political Economy 74 (2) (April), 132, doi: 10.1086/259131.

LeSage, J. P., 1999. The theory and practice of spatial econometrics. Available at: www.spatial-econometrics.com/html/sbook.pdf.

Lutzenhiser, M., Netusil, N. R., 2001. The effect of open spaces on a home's sales price. Contemporary Economic Policy 19 (3), 291–298.

Mahan, B. L., Polasky, S., Adams, R. M., 2000. Valuing urban wetlands: a property price approach. Land Economics 76 (1), 100–113.

Nelson, J. P., 1982. Highway noise and property values: a survey of recent evidence. Journal of Transport Economics and Policy 16 (2) (May), 117–138.

Palmquist, R. B., 1984. Estimating the demand for the characteristics of housing. Review of Economics and Statistics 66 (3) (August), 394–404.

Palmquist, R. B., 1992. Valuing localized externalities. Journal of Urban Economics 31 (1) (January), 59–68 .

Provencher, B., Sarakinos, H., Meyer, T., 2008. Does small dam removal affect local property values? an empirical analysis. Contemporary Economic Policy 26 (2) (April), 187–197.

Ridker, R. G., Henning, J. A., 1967. The determinants of residential property values with special reference to air pollution. Review of Economics and Statistics 49 (May), 246–57, doi: 10.1016/0094.1190(92)90032-G.

Rosen, S., 1974. Hedonic prices and implicit markets: product differentiation in pure competition. Journal of Political Economy 82 (1), 34–55.

Sander, H. A., Polasky, S., 2009. The value of views and open space: estimates from a hedonic pricing model for Ramsey County, Minnesota, USA. Land Use Policy 26, 837–845.

Shultz, S. D., King, D. A., 2001. The use of census data for hedonic price estimates of open-space amenities and land use. Journal of Real Estate Finance and Economics 22 (2/3), 239–252.

CHAPTER 4

HYPOTHESIS TESTING OF HEDONIC PRICE PARAMETERS WITH BOTH INPUT AND OUTPUT: AN APPLICATION IN CHENGDU, CHINA

Abstract

This study focuses on the profit maximization behavior of the property developer, which is the key to linking the factor market (i.e., the land market) and the commodity market (i.e., the housing market) together. With the true valuation of land derived in Chapter 2, we posit a profit function that considers all costs in the property development activity, including both the land cost and the non-land costs. A set of partial derivatives of the profit function with respect to various characteristics gives us the relationship between the marginal valuations in the land (factor) and housing (commodity) markets. We test the theoretical relationship using a spatial full information maximum likelihood (SFIML), which considers the land market, the housing market, and the profit maximization behavior of the property developer together in a joint likelihood function. Finally, we conduct hypothesis tests in both of these two scenarios to examine the validity of our linked-markets assumption in the hedonic price estimation.

1. Introduction

In this study, we focus on the profit maximization behavior of the property developer, which is the key to linking the factor market (i.e., the land market) and the commodity market (i.e., the housing market) together. The property developer is assumed to earn a positive profit arising from the premium in the English auction by which land is sold in China. We posit a profit function that considers all costs of production, including the land cost and the non-land costs.

A set of partial derivatives of the profit function with respect to various characteristics gives us the relationship between the marginal valuations in the land and housing markets, providing an opportunity to impose and test constraints on the estimated parameters. We use a spatial full information maximum likelihood (SFIML), which considers the land market, the housing market, and the profit maximization behavior of the property developer together in a joint likelihood function. We then use the results from the corresponding separate estimation in the housing market as a constraint on the SFIML parameters.

The paper proceeds as follows: after reviewing the related literature, we derive a theoretical model based on the profit maximization behavior of the property developer with respect to the various characteristics of land and housing. We then derive two empirical models: a separate estimation model and a joint estimation model. Hypothesis tests are conducted in both of these two scenarios.

2. Literature Review

Rosen's (1974) study is often credited as the seminal work in hedonic theory.

Given the focus of this paper, we draw from two strands of the related literature: hedonic theory with both inputs and outputs, and theoretical foundations of spatial urban models related to local public goods. In the first strand of literature, Heal (2001) and Palmquist (1989) have discussed the supply side of hedonic theory. Heal (2001) extends the model by introducing imperfect competition. Heal states that in a competitive market for a private good, willingness to pay is not captured by sellers, but rather is dissipated by competition and remains with the buyers as consumers surplus. He suggests that a separate market for the public good would not reach an efficient level of provision because of the classical free rider problem. A profit maximizing developer would provide a local public good at the economically optimal level, assuming the developer balances the tradeoff between development and conservation of open space. Obviously, the trade-off shows that more development means more houses to sell while more open space amenities may increase the per unit value of the houses. The household's willingness to pay for the change of open-space amenity is then derived from the indirect utility function which is the solution of the classical utility maximization problem.

The core argument of Heal's study is that a market need not automatically and efficiently provide local public goods. However, when the developer considers its profit maximization problem, it is possible that it could efficiently provide local public good. This would only occur when the developer is able to develop the locally recognized environmental asset exclusively. Heal considers this as a special case of first-order price discrimination, which can exist only when the utility function is strictly concave, and the cost function is strictly convex. In addition, the developer must have sole development rights. Housing prices are positively affected by local environmental quality; Heal argues that if the developer is aware of that, it could treat the provision of environmental assets as part of its profit-maximization problem.

While Rosen (1974) focuses on the different characteristics of the output, Palmquist (1989) extends the theory by considering the differentiated factors of production, with a focus on land. Palmquist uses land as an example of differentiated production inputs. He assumes that the rental price of land, R, depends on certain characteristics associated with the land, which are denoted by z_i, $i=1,2,..., n$, resulting in $R = R(z_1,z_2,...,z_n)$. Palmquist assumes this differentiated factor is provided by a landowner and is then used in production activities. The producer's problem is to maximize $\pi^{DV} = \sum_j^m p_j \times x_j$, subject to $g(x,z,\alpha) = 0$ and $\pi^{DV} \geq 0$, where π^{DV} is the producer's variable profit, vector p consists of the prices of the outputs and non-land inputs, and $g(x,z,\alpha) = 0$ is the implicit production function (vector x consists of outputs and inputs that could be positive or negative). The non-land input demand function is given by $x = x(p,z,\alpha)$, and hence the producer's bid function for a land parcel is denoted as $\theta(z,p,\pi^D,\alpha) = \pi^{*DV}(p,z,\alpha) - \pi^D$, where π^D is the desired profit level.

The supply side is also similar to the Rosen's (1974) model, but the characteristics are separated according to whether the landowner is able to alter them. Palmquist separates the characteristics vector z into two sub-vectors, $\hat{z} = (z_1, z_2, \ldots, z_i)$ and $\tilde{z} = (z_{i+1}, z_{i+2}, \ldots, z_n)$. $\hat{z}$ are characteristics exogenous to the landowner and the components of $\tilde{z}$ are within its control. The problem for the landowner, therefore, is to maximize $\pi^S = R(\tilde{z}; \hat{z}) - C(\tilde{z}; \hat{z}, r, \boldsymbol{\beta})$, subject to $\pi^S \geq 0$, where π^S represents the profit of the landowner, $R(\cdot)$ is the land rental price, $C(\cdot)$ is some cost function, r is a vector of input prices, and $\boldsymbol{\beta}$ is a vector of technology parameters.[45] Therefore, the offer function could be given as $\phi(\hat{z}, \tilde{z}, \pi^{S\prime}, r, \boldsymbol{\beta}) = \pi^{S\prime} + C(\hat{z}, \tilde{z}, r, \boldsymbol{\beta})$, where $\pi^{S\prime}$ is the desired profit level of the landowner. Finally, the combination of the bid function and the offer function defines the hedonic equilibrium function as the envelope of the tangent points, which is $P = P(\hat{z}, \tilde{z}, r, \boldsymbol{\beta})$.

In the second strand of the literature, Santerre (1985) and Wu (2006) discuss the theoretical foundation of spatial urban models that are related to local public goods. Local public goods can affect nearby property values to a significant degree. Santerre (1985) assumes that the median voter chooses between public and private inputs when producing the final output of housing services. The voter's preference can affect the local government's supply of local public goods. Santerre considers a monocentric metropolitan area composed of a number of jurisdictions. All economic activities take place in the central business district (CBD), and each ring located x

45 By implication, besides the common input price and technology level, the land owner's cost depends on its "endowment" $\hat{z}$, as well as how the land owner wants to "shape" $\tilde{z}$.

miles from the CBD corresponds to a different jurisdiction. Voters choose housing service and a composite private good to maximize their utility. In their budget constraint, the rental price of land, rental price of capital, the price of the composite private good, property tax share, money cost of travel are assumed to be functions of distance to the CBD. In addition, a congestion variable is introduced to represent the impure nature of the local public good. The solution to the utility maximization problem implicitly yields the demand for local public good.

Santerre uses data which are composed of 110 municipalities from the 11 SMSAs in Connecticut representing a broad range of local public goods, including local public education, fire, health, highway, library, parks and recreation, police, and sanitation service. The estimation results strongly suggest that distance from the CBD influences the composition of local public goods, and the findings show that demand for local public fire protection, parks and recreation, police, and sanitation services are negatively related to distance from CBD. In addition, Santerre estimates several substitution effects between residential land and different types of local public goods.

Wu (2006) develops an economic foundation to analyze urban development patterns and their relations to community characteristics. He divides amenities into two categories: one is exogenous amenities, which are defined as major geographic features such as rivers, scenery hills, and oceans, or by the idiosyncratic history of development; the other category is endogenous amenities, which are defined as local public services along with the location patterns of different income groups.[46] Wu's basic assumptions include: a central business district (CBD) which is exogenously determined, households with identical incomes and preferences, and travel cost, which is a function of the distance between residence and CBD. Amenities are assumed to vary over the urban landscape, which is setup in a Cartesian coordinate plane centered on the CBD.

Households are assumed to choose quantity of housing and composite goods, as well as the location coordinates. The indirect utility function is fixed at an exogenous baseline value, after which the household's bid-rent function is identified implicitly from the indirect utility function. Wu also considers the supply side of the housing market. A competitive residential development industry is assumed to have constant-returns-to-scale production technology. The developers choose a development density (defined as square feet of housing per acre of land) to maximize their profit, upon a location specific land rent. The zero profit condition yields the bid rent function for the developers.

Wu considers two types of landscapes: one is a line-featured amenity source such as a river or ocean shore; the other is an area-featured amenity sources such as a lake or a park. He proposes an exponential functional form to calculate the weighted summation of the total level of amenity at each residential site to every amenity source. The bigger the size of the source, and the closer the distance to it, the higher

46 For example, the Forbidden City lies in the center of the city of Beijing.

the total amenity level. With a Cobb-Douglas utility function, Wu numerically simulates the city's land use pattern with different settings of open space amenities. His simulation results show that while preserving land for open space removes some land from the path of land use, it may create incentives for even more development.

While most hedonic studies focus solely on the consumer side of the market, the critical role of the property developer has long been ignored. In fact, it is the property developer that links the land and housing markets together. Many studies just study these two markets separately. Although the study of Palmquist (1989) and Wu (2006), among others, have shed light on the theoretical link between the factor and commodity markets, no study that we are aware of has conducted hypothesis tests to empirically examine this linkage in hedonic price estimation.

3. Theoretical Model

While Wu's (2006) work appears to be the standard approach for spatial urban models related to local public goods, it does not focus on the role played by the property developer as a link between the factor (i.e., land) and commodity markets (i.e., the housing market). Although Palmquist (1989) points out the importance of combining the input and output together with respect to the differentiated characteristics, he does not demonstrate how to empirically test such a link. Our goal in this section, therefore, is to derive a theoretical model of land and housing markets that is empirically testable.

The second and third chapters have cast some light on this issue. The second chapter studies a Chinese regional land market. Given the fact that the local government plays the role of monopolistic land seller and English auction is used in the land transaction, the key feature of this chapter is to isolate the offer function of the seller (i.e., the local government) and the bid function of the buyers (i.e., the property developers), respectively. The auction premium (i.e., the difference between the transaction price and the asking price proposed by the local government) was used to derive the distribution of the buyer's valuation and hence the market participants' true valuations on land.

In addition, the third chapter studies the hedonic valuation for urban housing with both spatial and project heterogeneity in the same regional market. It uses housing retail sales data of 6 housing projects with hundreds of housing units each. Given the fact that housing development in China is relatively dense, the third chapter, therefore, utilizes a three-dimensional spatial weight matrix to control for spatial autocorrelation in the estimation of the hedonic price function. The reason why we build our model upon the second and third chapters is that these chapters provide separate hedonic estimations of the input (i.e., land) and output (i.e., housing) in the same market, i.e., Chengdu, China. If we wish to adequately link hedonic models for inputs and outputs together, we have to discuss them in the same market; otherwise the estimation results are neither consistent nor comparable.

Henceforth, we assume that the derived true valuation of the developer is θ_L, and the equilibrium housing retail sales price is P_H. Both θ_L and P_H are measured in ￥/m^2. Following the second and third chapters, we assume that θ_L and P_H are both functions of various housing and project characteristics. Thus, we write θ_L and P_H in the following manner:

$$\theta_L = \theta_L(PR, X_1, X_2) \tag{4.1}$$

$$P_H = P_H(PR, H, X_1, X_3) \tag{4.2}$$

where PR is the Plot Ratio (a density ratio that is the total floor area of the construction divided by the area of the associated land upon which the housing project is located), H is floor area of the housing unit, X_1 are characteristics common in the valuation of both land and housing, X_2 and X_3 are characteristics that are distinct in the valuation of land and housing, respectively. PR and X_1 are common attributes in the valuation of both land and housing; to link the input and output markets, we need to establish the theoretical relationship between the parameters for PR and X_1 in both markets.

In order to form the property developer's profit maximization problem, we need to introduce a cost function. Besides the expenditure on land, we consider a "development cost," which includes the design, construction, administrative, and tax costs, etc. In order to differentiate this development cost from the land cost, we call it "non-land costs," denote by C^{NL} as,

$$C^{NL} = C^{NL}(PR, H, X_1, X_2, X_3) \tag{4.3}$$

We consider the property developer's profit maximization problem at the housing-unit (apartment) level. At the time the land parcel is purchased, the developer builds several housing units at the same time, if attempting to maximize profit over all the housing units built on that land parcel. We evaluate the profit maximization problem at the housing unit level in order to better match the retail sales market data for apartments. Therefore, for each housing unit *i*, we can define the property developer's profit as:

$$\Pi_i = [P_{Hi}(PR,H_i,X_1,X_3) \times H_i] - [P_L(PR,X_1,X_2) \times L_{Hi}] - C_i^{NL}(PR,H_i,X_1,X_2,X_3) \tag{4.4}$$

where P_L is a per unit actual land sales price, and L_{Hi} is the amount of land associated with housing unit *i*. Recall that the Plot Ratio, *PR*, is the ratio of total floor area that could be built on the given land parcel divided by the area of that land. Thus, given a land parcel with area *L*, the total floor area that could be built on the land parcel is simply $L \times PR$. Since the share of the floor area of the target housing unit is $H_i / (L \times PR)$, the share of the land area associated with the housing unit is then given by:[47]

47 The property developer does not necessarily sell all housing units in its housing project at the same time. Normally, it develops a housing project in "batches." The developer could build housing units in some part of the land parcel, and sell them; then develop another part of the land parcel, etc. Assuming the developer sells *n* units of housing at a given time, the total floor area of these housing units is $\sum_{i=1}^{n} H_i$. Obviously, the ratio of this batch of housing units to the total available units is $\sum_{i=1}^{n} H_i / (L \times PR)$. Considering the share of the target housing unit in this batch of housing units, $H_i / \sum_{i=1}^{n} H_i$, we could write the expression for L_{Hi} as:

$$L_{Hi} = L \times [H_i / (L \times PR)] = H_i / PR \qquad (4.5)$$

Eq. (4.5) is straight forward, since the amount of land associated with the housing unit is simply the unit's floor area divided by the Plot Ratio.

In the second chapter, we have calculated the developer's derived valuation θ_L with an assumption of zero profit, which means that θ_L is the maximum amount that the developer would bid in the English auction. However, if the price actually paid for land (P_L) is less than θ_L, the property developer clearly anticipates a positive profit in its property development activity.

Assuming that all factors of production other than land are paid according to the value of their marginal product, the anticipated profit from development is θ_L - P_L, thus for each housing unit, the anticipated profit is $(\theta_L - P_L) \times L_{Hi}$. This anticipated profit arises from incentive incompatibility problem in the English auction, since the winner only needs to pay the amount at which the second highest bidder quits plus the last increment in the auction. Given Eq. (4.5), we can rewrite Eq. (4.4) as follows:

$$[\theta_L(PR,X_1,X_2) - P_L] \times H_i / PR$$

$$= P_{Hi}(PR,H,X_1,X_3) \times H_i - P_L(PR,X_1,X_2) \times H_i / PR - C_i^{NL}(PR,H,X_1,X_2,X_3) \qquad (4.6)$$

In Eq. (4.6), the left hand side is the anticipated profit associated with development of the target housing unit. On the right hand side, $P_{Hi}(PR,H_i,X_1,X_3) \times H_i$ is the revenue of selling housing unit i, $\theta_L(PR,X_1,X_2) \times (H_i / PR)$ is the allocated

---

$L_{Hi} = L \times [\sum_{i=1}^{n} \mathrm{H_i} / (L \times PR)] \times (H_i / \sum_{i=1}^{n} \mathrm{H_i}) = H_i / PR$, which is exactly the same as that in Eq. (4.5).

land cost for housing unit *i*, and $C_i^{NL}(PR,H_i,X_1,X_2,X_3)$ is the non-land costs associated with development of housing unit *i* besides the expenditure on land. Although we have no preliminary assumption about the functional form of the non-land costs C_i^{NL}, apparently it could be solved immediately from a transformation of Eq. (4.6) as:

$$C_i^{NL}(PR,H_i,X_1,X_2,X_3) = P_{Hi}(PR,H_i,X_1,X_3) \times H_i - \theta_L(PR,X_1,X_2) \times (H_i / PR) \quad (4.7)$$

Since $PR>0$ and $H_i >0$, we can rearrange Eq. (4.6) in the following manner:

$$\theta_L(PR,X_1,X_2) = P_{Hi}(PR,H,X_1,X_3) \times PR - C_i^{NL}(PR,H,X_1,X_2,X_3) \times PR / H_i \quad (4.8)$$

The property developer is assumed to be a profit maximizer, therefore, we could consider it to be maximizing its profit over the various characteristics. Given our research problem in this paper, we are only interested in those common characteristics, i.e., PR and X_1. We then could differentiate Eq. (4.8) with respect to PR and X_1 as follows:

$$\partial\theta_L(PR,X_1,X_2)/\partial PR = \underbrace{[\partial P_{Hi}(PR,H,X_1,X_3)/\partial PR] \times PR + P_{Hi}(PR,H,X_1,X_3)}_{\text{marginal revenue with respect to } PR}$$

$$- \underbrace{[\partial C_i^{NL}(PR,H,X_1,X_2,X_3)/\partial PR] \times PR / H_i - C_i^{NL}(PR,H,X_1,X_2,X_3) / H_i}_{\text{marginal cost with respect to } PR} \quad (4.9)$$

$$\partial\theta_L(PR,X_1,X_2)/\partial X_1 = \underbrace{[\partial P_{Hi}(PR,H,X_1,X_3)/\partial X_1] \times PR}_{\text{marginal revenue with respect to } X_1}$$

$$- \underbrace{[\partial C_i^{NL}(PR,H,X_1,X_2,X_3)/\partial X_1] \times PR / H_i}_{\text{marginal cost with respect to } X_1} \quad (4.10)$$

Eqs. (4.9) and (4.10) are what we need to link the estimation parameters for the hedonic price functions both in the land and housing markets. They are not new to us if we recall the case of monopolistic seller in the standard production theory. In our case, though, the market is assumed to be competitive, since the prices are functions of the characteristics, the terms in the big bracket are nothing but the standard "marginal revenue minus marginal cost."

Marginal revenue in Eq. (4.9) is composed of the sales price of the marginal housing unit, plus the change in the price of all housing units in response to a change in housing density. Marginal costs are composed of per unit area non-land costs, plus the change in per unit non-land costs associated with a change in housing density. Eq. (4.10) is interpreted similarly for characteristics X_1.

4. Data

Our study uses the same data sets used in the second and third chapters. The second chapter uses a land wholesale data set in Chengdu city, China. Chengdu is one of largest inland cities in mainland China. As the capital city of Sichuan province, it lies in the southwestern part of China, which is about 1500 kilometers southwest of Beijing. With nearly 13 million official residents, the shape of Chengdu is a standard monocentric city, the core metropolitan area of which consists of four concentric ring roads and several radius roads.

Three hundred and fifty effective land transactions for residential development were recorded between January 2004 and October 2009, most of which lie between the second and the third ring road (34%), and are then manually geo-coded into GIS coordinates. In the land data set, we have information on transaction date, type of English auction, the size of the parcel, asking price by the local government, and the actual transaction price.[48] We also have information on development regulations, such as the Plot Ratio. In addition, for each of the land parcel, we have information on calculated proximity and aggregate level to different sources of open space amenities and local infrastructure, such as parks, rivers, subway stations, roads, and public hospitals.[49,50] Descriptive statistics for this data set are contained in Tables 2.1 and 2.2 in the second chapter.

The housing retail sales data set used in the third chapter consists of six housing projects, four of which are located around the third ring road. This data set includes 1,268 observations in total, with an average of 211 housing units in each housing project. This data set includes information on both the project and housing-unit attributes. Plot Ratio, Subway Station Proximity, Park Proximity, Hospital Aggregate, and Road Aggregate are included in the former category. In the latter category, there are "*view*" dummies, such as whether the housing unit has a view to an open-space

48 The only difference between the two types of auction is the length of time period that the potential buyers could bid.
49 Proximity is calculated as the distance from the centroid of a land parcel to the centroid of a source of open space amenity or local infrastructure using the Haversine Formula.
50 Aggregate level is calculated as weight aggregate of the "service level" from each source of open space amenity or local infrastructure, using the inverse distance from centroid to centroid as the weights. The service level varies depending on the type of the source, such as the size of the park, the number of available beds in the hospital, etc.

amenity sources either within or outside the housing project, whether the unit is located adjacent to a major road, as well as the direction the housing units face.[51] In addition, housing sales price, the floor number, area of the housing unit, distance to major open-space amenity within each project, and transaction date are also included in the data set as well. Descriptive statistics for this data set are included in Tables 3.1 and 3.2 in the third chapter.

In this study, we need to combine the land wholesale data set and the housing retail sales data set together. Unfortunately, we can only identify 4 of the 6 housing projects as the raw land parcels in the land data set with 350 observations (see Fig. 4.1). This difficulty limits us to only 3 characteristics that are common in the estimation of both the land and housing markets. We therefore choose those factors that are both rich in economic meaning and statistically significant. Three common factors include Plot Ratio, proximity to subway station, and aggregate service level of urban road. Using the notation described in section 3, Plot Ratio is "*PR*," proximity to subway station and aggregate road service level are "X_1." All other explanatory variables in the second chapter are treated as "X_2," and those in the third chapter are treated as "X_3."

For the four housing projects included in this study, we have 851 observations,

51 In the Chinese culture having the major rooms facing south (such as living room, main bedroom, etc.) is considered to be the most preferable direction for living.

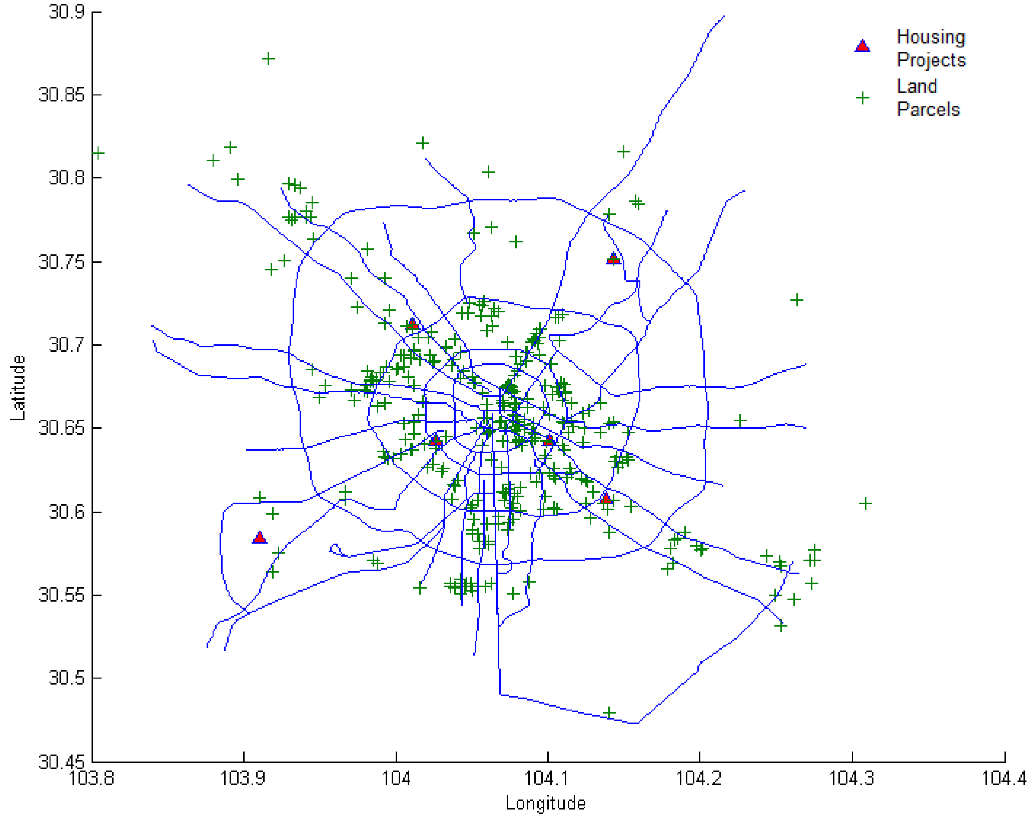

Fig. 4.1. Spatial Distribution of Land Parcels and Housing Projects

roughly 213 housing units per housing project on average over 46% of the housing units have a view to major open space amenities within the housing projects, 40% of the housing units have a view to some source of open space amenities outside the projects, while 34% of the housing units situate close to a major urban road. In regard to the direction of the major rooms in the housing units, 17% are facing north, 8% northeast, 19% southeast, 7% south, 2% southwest, 4% west, and 30% northwest. No observations faced east. In addition, about 75% of the housing units are purchased via mortgage, while 25% are paid in cash. Descriptive statistics are reported in Table 4.1.

5. Empirical Models

5.1. Separate Estimation Model

Following the semi-log functional form used in the estimations of Chapters 2 and 3, we assume that both θ_L and P_H have exponential forms. Thus, we can rewrite θ_L and P_H as follows:

$$\theta_L = exp(\boldsymbol{\beta}_{L\text{-}PR}\, PR + \boldsymbol{\beta}_{L\text{-}X1}\, X_1 + \boldsymbol{\beta}_{L\text{-}X2}\, X_2) \tag{4.11}$$

$$P_H = exp(\boldsymbol{\beta}_{H\text{-}PR}\, PR + \boldsymbol{\beta}_{H\text{-}H}\, H + \boldsymbol{\beta}_{H\text{-}X1}\, X_1 + \boldsymbol{\beta}_{H\text{-}X3}\, X_3) \tag{4.12}$$

where $\boldsymbol{\beta}_{L\text{-}PR}$, $\boldsymbol{\beta}_{L\text{-}X1}$ and $\boldsymbol{\beta}_{L\text{-}X2}$ are the parameters for PR, X_1 and X_2 estimated in the second chapter, and $\boldsymbol{\beta}_{H\text{-}PR}$, $\boldsymbol{\beta}_{H\text{-}H}$, $\boldsymbol{\beta}_{H\text{-}X1}$ and $\boldsymbol{\beta}_{H\text{-}X3}$ are the parameters for PR, H, X_1 and X_3 estimated in the third chapter. In addition, we assume that the non-land costs C^{NL}, has a semi-log functional form with respect to various characteristics as well. Thus, we have:

$$C^{NL} = exp(\boldsymbol{\beta}_{C\text{-}PR}\, PR + \boldsymbol{\beta}_{C\text{-}H}\, H + \boldsymbol{\beta}_{C\text{-}X1}\, X_1 + \boldsymbol{\beta}_{C\text{-}X2}\, X_2 + \boldsymbol{\beta}_{C\text{-}X3}\, X_3) \tag{4.13}$$

However, since we only have data on 4 housing projects that could be identified as the raw land parcels in the land wholesale data set, we are unable to implement the regression on the full set of explanatory variables in Eq. (4.13). We therefore have to eliminate X_2 from Eq. (4.13), and hence we end up with the following functional form for C^{NL}:

Table 4.1
Descriptive Statistics for the 4-project Data Set

	Mean	Median	Std. Dev.
Project attributes			
Plot Ratio	4.363	4.780	0.731
Subway station proximity (m)	3000.556	1474.832	2419.188
Road level	8.831	10.295	2.207
Land unit sales price (￥/ m^2)	7393.974	10574.990	4011.584
Deflated land unit sales price (￥/ m^2)	5843.329	8254.638	3112.034
True land valuation with deflation only (￥/ m^2)	8063.476	9024.962	2269.026
True land valuation with only dummy (￥/ m^2)	9807.723	11273.550	2959.034
True land valuation with both deflation and dummy (￥/ m^2)	7557.051	8758.900	2399.368
Housing unit attributes			
Distance to major open-space amenity within each project (m)	212.842	99.499	257.032
Housing unit area (m^2)	92.469	85.730	1418.518
Housing unit sales price (￥/ m^2)	6087.727	5875.110	0.003
Deflated housing unit sales price (￥/ m^2)	4519.693	4356.167	1033.613
Floor in which the unit situates	12.800	11.000	7.750

$$C^{NL} = exp(\boldsymbol{\beta}_{C\text{-}PR}\, PR + \boldsymbol{\beta}_{C\text{-}H}\, H + \boldsymbol{\beta}_{C\text{-}X1}\, X_1 + \boldsymbol{\beta}_{C\text{-}X3}\, X_3) \tag{4.14}$$

The corresponding regression equation in the log form is as follows:

$$log(C^{NL}) = \alpha_C + \boldsymbol{\beta}_{C\text{-}PR}\, PR + \boldsymbol{\beta}_{C\text{-}H}\, H + \boldsymbol{\beta}_{C\text{-}X1}\, X_1 + \boldsymbol{\beta}_{C\text{-}X3}\, X_3 + e_C \tag{4.15}$$

where α_C and e_C, respectively, are the corresponding intercept term and random error component in the regression. As pointed out in the second and third chapters, time is an important factor that could affect the estimation results for the cross sectional data covering a long period of time. We therefore have three choices in dealing with the time effect: deflation only by certain price index, either housing price index (HPI) or consumption price index (CPI), only time dummies, as well as a combination of these two approaches. It is completely an empirical issue to determine which one to be the best. In this study we implement each of these three approaches, and then apply the corresponding version of θ_L and P_H to Eqs. (4.9) and (4.10), respectively. The results of the estimation of Eq. (4.15) are reported in Table 4.2 with deflation only by a monthly housing price index, in Table 4.3 with only monthly dummies but no deflation, and in Table 4.4 with both deflation and dummy.[52,53]

Although the regression of the non-land costs C^{NL}, is an intermediate step for us to form the constraint that is used later in our analysis, it is worth spending some time discussing the coefficients values in Tables 4.2 to 4.4. The negative sign on the coefficient for Plot Ratio is as expected, since the higher the housing project's density, the more the developer could build. Thus, it is simply an economy of scale that the non-land costs for each housing unit would be lower. The positive sign of Road aggregate is consistent with common sense, since a location that has denser aggregate road service level is commonly more commercialized. Construction in such location incurs higher costs than elsewhere. For example, comparing a housing project in downtown with an identical housing project in a suburban area, it would be more

52 This housing price index of Chengdu is reported monthly by an authority called the National Development and Reform Commission.
53 We set the monthly time trend variable, starting with January 2004 equal to one and ending with October 2009 equal to 70.

Table 4.2
OLS Estimation of Unit Non-land Costs, C^{NL} (￥/ m^2), Using HPI Deflation

Variable	Coefficient	t-statistic	p-value
Intercept	11.884	97.534	0.000
Project-Attribute Variables ("PR" and "X_1")			
Plot Ratio	-0.441	-13.263	0.000
ln(Subway station proximity)	0.002	0.181	0.856
Road aggregate	0.175	56.215	0.000
Housing-Unit-Attribute Variables ("H" and "X_3")			
Floor	0.001	3.926	0.003
Inner view (1=yes)	0.183	20.188	0.000
Outer view (1=yes)	0.028	3.245	0.000
Close to street (1=yes)	0.131	13.940	0.000
North (1=yes)	-0.067	-6.321	0.000
North East (1=yes)	-0.132	-9.038	0.000
South East (1=yes)	-0.023	-2.646	0.021
South (1=yes)	-0.078	-5.283	0.000
South West (1=yes)	-0.031	-2.591	0.005
West (1=yes)	-0.074	-4.330	0.008
Distance to inner source	-0.000	-0.819	0.132
Housing unit area	0.010	68.374	0.000
Pay in cash (1=yes)	-0.023	-3.426	0.000
Adjusted R-square		0.974	
sigma^2		0.006	

Dependent variable: ln(Per unit housing cost with deflation only)

costly to transport construction waste out of the city, or there may be additional fees charged for the noise and dust associated with the construction in the downtown area.

Now, substitute Eqs. (4.11) and (4.12), along with (4.14), into Eqs. (4.9) and (4.10), we have:

$$exp(\boldsymbol{\beta}_{L\text{-}PR}\, PR + \boldsymbol{\beta}_{L\text{-}X1}\, X_1 + \boldsymbol{\beta}_{L\text{-}X2}\, X_2) \times \boldsymbol{\beta}_{L\text{-}PR}$$

Table 4.3
OLS Estimation of Unit Non-land Costs, C^{NL} (¥/ m^2), Using Monthly Time Dummy

Variable	Coefficient	t-statistic	p-value
Intercept	12.461	102.011	0.000
Project-Attribute Variables ("PR" and "X_1")			
Plot Ratio	-0.429	-11.336	0.000
ln(Subway station proximity)	-0.040	-3.244	0.000
Road aggregate	0.165	28.619	0.000
Housing-Unit-Attribute Variables ("H" and "X_3")			
Floor	0.002	4.390	0.000
Inner view (1=yes)	0.178	19.541	0.000
Outer view (1=yes)	0.029	3.305	0.000
Close to street (1=yes)	0.129	13.740	0.000
North (1=yes)	-0.064	-6.043	0.000
North East (1=yes)	-0.135	-9.185	0.000
South East (1=yes)	-0.024	-2.680	0.012
South (1=yes)	-0.077	-5.208	0.000
South West (1=yes)	-0.032	-2.659	0.005
West (1=yes)	-0.074	-4.314	0.002
Distance to inner source	-0.000	-0.912	0.161
Housing unit area	0.010	66.907	0.000
Pay in cash (1=yes)	-0.021	-3.160	0.000
Time trend	0.002	1.810	0.095
Adjusted R-square		0.972	
sigma^2		0.006	

Dependent variable: ln(Per unit housing cost with only dummy)

$$
\begin{aligned}
&= exp(\boldsymbol{\beta}_{H\text{-}PR}\, PR + \boldsymbol{\beta}_{H\text{-}H}\, H_i + \boldsymbol{\beta}_{H\text{-}X1}\, X_1 + \boldsymbol{\beta}_{H\text{-}X3}\, X_3) \times \boldsymbol{\beta}_{H\text{-}PR}\, PR \\
&\quad + exp(\boldsymbol{\beta}_{H\text{-}PR}\, PR + \boldsymbol{\beta}_{H\text{-}H}\, H_i + \boldsymbol{\beta}_{H\text{-}X1}\, X_1 + \boldsymbol{\beta}_{H\text{-}X3}\, X_3) \\
&\quad - exp(\boldsymbol{\beta}_{C\text{-}PR}\, PR + \boldsymbol{\beta}_{C\text{-}H}\, H_i + \boldsymbol{\beta}_{C\text{-}X1}\, X_1 + \boldsymbol{\beta}_{C\text{-}X3}\, X_3) \times \boldsymbol{\beta}_{C\text{-}PR}\, PR / H_i \\
&\quad - exp(\boldsymbol{\beta}_{C\text{-}PR}\, PR + \boldsymbol{\beta}_{C\text{-}H}\, H_i + \boldsymbol{\beta}_{C\text{-}X1}\, X_1 + \boldsymbol{\beta}_{C\text{-}X3}\, X_3) / H_i \qquad (4.16)
\end{aligned}
$$

Table 4.4
OLS Estimation of Unit Non-land Costs, C^{NL} (￥/ m^2), Using Both HPI Deflation and Monthly Dummy

Variable	Coefficient	t-statistic	p-value
Intercept	11.870	97.366	0.000
Project-Attribute Variables ("PR" and "X_1")			
Plot Ratio	-0.474	-12.531	0.000
ln(Subway station proximity)	-0.003	-0.234	0.815
Road aggregate	0.184	31.931	0.000
Housing-Unit-Attribute Variables ("H" and "X_3")			
Floor	0.002	3.964	0.000
Inner view (1=yes)	0.183	20.173	0.000
Outer view (1=yes)	0.028	3.256	0.001
Close to street (1=yes)	0.132	14.052	0.000
North (1=yes)	-0.066	-6.223	0.000
North East (1=yes)	-0.134	-9.197	0.000
South East (1=yes)	-0.024	-2.735	0.006
South (1=yes)	-0.077	-5.197	0.000
South West (1=yes)	-0.032	-2.694	0.007
West (1=yes)	-0.074	-4.353	0.001
Distance to inner source	-0.000	-0.859	0.391
Housing unit area	0.010	66.848	0.000
Pay in cash (1=yes)	-0.022	-3.398	0.001
Time trend	0.002	1.964	0.050
Adjusted R-square		0.974	
sigma^2		0.006	

Dependent variable: ln(Per unit housing cost with both deflation and dummy)

$$exp(\boldsymbol{\beta}_{L\text{-}PR}\, PR + \boldsymbol{\beta}_{L\text{-}X1}\, X_1 + \boldsymbol{\beta}_{L\text{-}X2}\, X_2) \times \boldsymbol{\beta}_{L\text{-}X1}$$

$$= exp(\boldsymbol{\beta}_{H\text{-}PR}\, PR + \boldsymbol{\beta}_{H\text{-}H}\, H_i + \boldsymbol{\beta}_{H\text{-}X1}\, X_1 + \boldsymbol{\beta}_{H\text{-}X3}\, X_3) \times \boldsymbol{\beta}_{H\text{-}X1}\, PR$$

$$- exp(\boldsymbol{\beta}_{C\text{-}PR}\, PR + \boldsymbol{\beta}_{C\text{-}H}\, H_i + \boldsymbol{\beta}_{C\text{-}X1}\, X_1 + \boldsymbol{\beta}_{C\text{-}X3}\, X_3) \times \boldsymbol{\beta}_{C\text{-}X1}\, PR / H_i \qquad (4.17)$$

After some rearrangement, Eqs. (4.16) and (4.17) can be rewritten as follows:

$$\theta_L \times \boldsymbol{\beta}_{L\text{-}PR} = P_{Hi} \times \boldsymbol{\beta}_{H\text{-}PR}\, PR + P_{Hi} - C_i^{NL} \times \boldsymbol{\beta}_{C\text{-}PR}\, PR / H_i - C_i^{NL} / H_i \qquad (4.18)$$

$$\theta_L \times \boldsymbol{\beta}_{L\text{-}X1} = P_{Hi} \times \boldsymbol{\beta}_{H\text{-}X1}\, PR - C_i^{NL} \times \boldsymbol{\beta}_{C\text{-}X1}\, PR / H_i \qquad (4.19)$$

Since Eqs. (4.18) and (4.19) hold at the per housing unit level, in order to use our data and implement the hypothesis test, we need to add these equations over all observations in our data set. We use j to denote the j^{th} housing project, for $j=1, 2, 3, 4$. We also assume that in each of these housing projects, there are a total of n_j housing units. Denoting each housing unit by i, noting the fact that θ_L and PR vary among different housing projects, and the estimation parameters (i.e., the $\boldsymbol{\beta}$s) are constant across all the housing units in the data, we therefore can add up Eqs. (4.18) and (4.19) in the following manner:

$$\boldsymbol{\beta}_{L\text{-}PR}\sum_{j=1}^{4}\sum_{i=1}^{nj}\theta_{Lj} = \boldsymbol{\beta}_{H\text{-}PR}\sum_{j=1}^{4}\sum_{i=1}^{nj}(P_{Hij}\times PR_j) + \sum_{j=1}^{4}\sum_{i=1}^{nj}P_{Hij} - \boldsymbol{\beta}_{C\text{-}PR}\sum_{j=1}^{4}\sum_{i=1}^{nj}(\frac{C_{ij}^{NL}\times PR_j}{H_{ij}}) - \sum_{j=1}^{4}\sum_{i=1}^{nj}(\frac{C_{ij}^{NL}}{H_{ij}}) \qquad (4.20)$$

$$\boldsymbol{\beta}_{L\text{-}X1}\sum_{j=1}^{4}\sum_{i=1}^{nj}\theta_{Lj} = \boldsymbol{\beta}_{H\text{-}X1}\sum_{j=1}^{4}\sum_{i=1}^{nj}(P_{Hij}\times PR_j) - \boldsymbol{\beta}_{C\text{-}X1}\sum_{j=1}^{4}\sum_{i=1}^{nj}(\frac{C_{ij}^{NL}\times PR_j}{H_{ij}}) \qquad (4.21)$$

Denoting $\overline{\beta_{\text{L-PR}}}$ and $\overline{\beta_{\text{L-X1}}}$ as the calculated constraints on the estimation parameters, Eqs. (4.20) and (4.21) can be rearranged as follows:

$$\overline{\beta_{\text{L-PR}}} = \{\boldsymbol{\beta}_{H\text{-}PR}\sum_{j=1}^{4}\sum_{i=1}^{nj}(P_{Hij}\times PR_j) + \sum_{j=1}^{4}\sum_{i=1}^{nj}P_{Hij} - \boldsymbol{\beta}_{C\text{-}PR}\sum_{j=1}^{4}\sum_{i=1}^{nj}(\frac{C_{ij}^{NL}\times PR_j}{H_{ij}}) - \sum_{j=1}^{4}\sum_{i=1}^{nj}(\frac{C_{ij}^{NL}}{H_{ij}})\} / (\sum_{j=1}^{4}\sum_{i=1}^{nj}\theta_{Lj}) \qquad (4.22)$$

$$\overline{\beta_{\text{L-X1}}} = \{\boldsymbol{\beta}_{H\text{-}X1}\sum_{j=1}^{4}\sum_{i=1}^{nj}(P_{Hij}\times PR_j) - \boldsymbol{\beta}_{C\text{-}X1}\sum_{j=1}^{4}\sum_{i=1}^{nj}(\frac{C_{ij}^{NL}\times PR_j}{H_{ij}})\} / (\sum_{j=1}^{4}\sum_{i=1}^{nj}\theta_{Lj}) \qquad (4.23)$$

Table 4.5
Summary of the Target Estimation Parameters in Land

	$\beta_{L\text{-}PR}$	$\beta_{L\text{-}Road\text{-}aggregate}$	$\beta_{L\text{-}Subway\text{-}proximity}$
Deflation only	0.132	0.045	-0.080
Only dummy	0.079	0.077	-0.040
Both deflation and dummy	0.077	0.076	-0.044

Source: The second chapter

Table 4.6
Summary of the Target Estimation Parameters in Housing

	$\beta_{H\text{-}PR}$	$\beta_{H\text{-}Road\text{-}aggregate}$	$\beta_{H\text{-}Subway\text{-}proximity}$
Deflation only	-0.301	-0.227	-0.595
Only dummy	-0.271	-0.200	-0.583
Both deflation and dummy	-0.299	-0.226	-0.595

Source: The third chapter

Eqs. (4.22) and (4.23) are what we need to conduct the hypothesis testing. We, therefore, setup the null hypothesis as follows:

$$\beta_{L\text{-}PR} = \overline{\beta_{\text{L-PR}}} \tag{4.24}$$

$$\beta_{L\text{-}X1} = \overline{\beta_{\text{L-X1}}} \tag{4.25}$$

In Eqs. (4.24) and (4.25), $\beta_{L\text{-}PR}$ and $\beta_{L\text{-}X1}$ are the estimation parameters in the second chapter, whereas $\overline{\beta_{\text{L-PR}}}$ and $\overline{\beta_{\text{L-X1}}}$ are calculated in this study in Eqs. (4.22) and (4.23), with $\beta_{H\text{-}PR}$ and $\beta_{H\text{-}X1}$ obtained from the third chapter. A summary of these target estimation parameters for land and housing from previous chapters are reported in Table 4.5 and Table 4.6, respectively. In addition, the calculated constraints for the target estimation parameters are reported in Table 4.7. We therefore conduct an F test for each set of the constraints. With deflation only by HPI, the calculated value of

Table 4.7
Calculated Constraints for the Target Estimation Parameters

	$\beta_{L\text{-}PR}$	$\beta_{L\text{-}Road\text{-}aggregate}$	$\beta_{L\text{-}Subway\text{-}proximity}$
Deflation only	0.2630	-1.2034	-2.0465
Only dummy	0.3701	-1.1649	-2.0179
Both deflation and dummy	0.3443	-1.2179	-2.0275

the F statistics is 2129.6; with only monthly time dummy, the calculated value of the F statistics is 2888.1; with both deflation and dummy, the calculated value of the F statistics is 3149.5. Given the critical value of 4.61 at 1% significance level or 3.00 at 5% significance level, we reject the null hypothesis of the valid constraints in each of these three cases.

This result should come with no surprise that, we have rejected the null hypothesis of equal parameters in the land estimation and in the derivation from the housing market. As Ellickson (1981) has pointed out, complex relationships in the hedonic price function often result in the varying estimation coefficients from neighborhood to neighborhood. In our case, although we have 350 land parcels scattered throughout the city, we only have 4 housing projects that can be identified as the raw land parcels. Hence, the representative power of these 4 housing projects could be limited or even somewhat biased with respect to the whole 350 land parcels.

5.2. Joint Estimation Model

As Ellickson (1981) has argued, the ability of hedonic theory to treat housing characteristics simultaneously is obscured by one's practical approach of discussing one attribute at a time, while treating others as fixed. In our case, characteristics from both the land market and the housing market, as well as the profit maximization behavior of the property developer, could have either positive or negative impacts on the housing sales price. Therefore, to link the land and housing markets together, it is necessary to conduct an overall examination of the effect of various characteristics on the housing price.

To our knowledge, few studies have been done to conduct hedonic analysis through joint estimation. One of the few exceptions is Al Refai (1994). Al Refai divides housing assets into land and structure, and uses an iterative three-stage least square (I3SLS) approach to simultaneously estimate land, structure, as well as the proportion of housing in total wealth. Given the spatial nature of our research problem, 3SLS approach is not appropriate to us (primarily for the efficiency issue). We therefore turn to another joint estimation method, i.e., the full information maximum likelihood (FIML) approach.

Discussion of FIML dates back to Chow (1968) and Eisenpress (1962), among others. Since then, FIML has been applied in a wide range of studies. We consider a joint estimation method that combines land market sales, housing market sales, and the non-land costs together, using FIML estimation. We therefore consider a joint probability density function $\boldsymbol{f}_{joint}(\theta_L,P_H,C^{NL})$, which can be rewritten via an application of Bayes's rule as follows:

$$\boldsymbol{f}_{joint}(\theta_L,P_H,C^{NL}) = \boldsymbol{f}_{\theta L}(\theta_L) \times \boldsymbol{f}_{PH}(P_H|\theta_L) \times \boldsymbol{f}_C(C^{NL}|\theta_L,P_H) \quad (4.26)$$

Land and housing market evaluations can be considered as two independent processes, however, the distribution of non-land costs C^{NL} is determined by the linkage of the land and housing markets, as shown in Eq. (4.7). Thus, we can rewrite Eq. (4.26) as:

$$\boldsymbol{f}_{joint}(\theta_L,P_H,C^{NL}) = \boldsymbol{f}_{\theta L}(\theta_L) \times \boldsymbol{f}_{PH}(P_H) \times \boldsymbol{f}_C(C^{NL}|\theta_L,P_H) \quad (4.27)$$

where, $\boldsymbol{f}_{\theta L}(\theta_L)$ and $\boldsymbol{f}_{PH}(P_H)$ are the probability density functions of the true land valuation and the housing retail sales price, respectively, and $\boldsymbol{f}_C(C^{NL}|\theta_L,P_H)$ is the conditional probability density function of non-land costs, which depends upon the true land valuation and the housing retail sales price. Eq. (4.7) describes a linear relationship among θ_L, P_H and C^{NL}, which is different from our previous models which used the log form of the key variables, i.e., $ln\theta_L$, lnP_H and lnC^{NL}. Therefore, for each of the three variables in our previous approach, Eq. (4.27) must be measured in its original form, not in log form.

Before we proceed to derive Eq. (4.27), it is worth spending some time discussing the structure of the equation's three components. From our previous discussion, we know that OLS estimation works fairly well for non-land costs. However, estimation of the housing retail sales price does best using the spatial error model (SEM), which is quite different from the OLS model. The standard SEM has a functional form as follows:

$$y = X \times \boldsymbol{\beta} + \eta,\ \eta = \boldsymbol{\lambda} \times W \times \eta + e \quad (4.28)$$

where $\boldsymbol{\lambda}$ is the spatial lag coefficient, and W is commonly known as the spatial weight matrix, which is typically constructed with element i, j as the inverse distance between location i and j. In this case, $y - X \times \boldsymbol{\beta}$ is no longer the actual disturbance term. Rather, $\eta = \boldsymbol{\lambda} \times W \times \eta + e$ is the true error term, which can be transformed as follows:

$$e = (I - \boldsymbol{\lambda} \times W) \times \eta \qquad (4.29)$$

Thus, η can be expressed as $(I - \boldsymbol{\lambda} \times W)^{-1} \times e$, and hence we have:

$$y = X \times \boldsymbol{\beta} + (I - \boldsymbol{\lambda} \times W)^{-1} \times e \qquad (4.30)$$

Now again consider Eq. (4.27). In the second chapter, the distribution of the true land valuation θ_L is transformed from the distribution of the disturbance term in a Tobit model estimation as follows:

$$f_{\theta L}(\theta_L) = f_{eL}(\theta_L - r - \boldsymbol{\beta}_{Tobit} X_L) \qquad (4.31)$$

where r is the asking price proposed by the land seller (i.e., the local government) in the English auction, $\boldsymbol{\beta}_{Tobit}$ is the vector of estimation parameters of the Tobit model, X_L is the set of explanatory variables in the land market, and $f_{eL}(\cdot)$ is the probability density function of the disturbance term in the Tobit estimation, which is distributed as $N(0, \sigma_L^2)$.

Transforming the distribution of e_L to the distribution of θ_L, we know that the mean of θ_L, $\boldsymbol{\mu}_L$, is $\boldsymbol{\beta}_{Tobit} X_L + r$, and the variance of θ_L, $\sigma_{\theta L}^2$, is simply σ_L^2. For the distribution of the housing retail sales price, by Eq. (4.30),we have:

$$P_H = X_H \times \boldsymbol{\beta}_H + (I - \boldsymbol{\lambda} \times W)^{-1} \times e_H \qquad (4.32)$$

where X_H is the set of explanatory variables in the housing market, $\boldsymbol{\beta}_H$ is the corresponding vector of housing estimation parameters, and e_H is the disturbance term, which is distributed as $N(0,\sigma_H^2)$. Thus, the disturbance term can be expressed as follows:

$$e_H = (I - \boldsymbol{\lambda} \times W) \times (P_H - X_H \times \boldsymbol{\beta}_H) \tag{4.33}$$

Transforming the distribution of e_H to the distribution of P_H, we know that the mean of P_H, $\boldsymbol{\mu}_H$, is $X_H \times \boldsymbol{\beta}_H$, but the variance of P_H, σ_{PH}^2, is a bit more complicated to derive.

From Eq. (4.32) we know that it is the term $(I - \boldsymbol{\lambda} \times W)^{-1} \times e_H$ that determine σ_{PH}^2. Denoting $P = (I - \boldsymbol{\lambda} \times W)^{-1}$, the i^{th} row of $P \times e_H$ is then expressed as $\sum_j P_{ij} \times e_{Hj}$, where j denotes the j^{th} element in each row of P. P_H is a vector, hence the i^{th} element of its variance is $\sum_j P_{ij}^2 \times \sigma_H^2$. According to the standard homoscedasticity assumption, we can add up all the terms and make average, and then we have:

$$\sigma_{PH}^2 = \sigma_H^2 \times \sum_i \sum_j P_{ij}^2 / n \tag{4.34}$$

where n denotes the number of observations in the data set. Using the property of trace operation, "$trace\ (P' \times P) = \sum_i \sum_j P_{ij}^2$," Eq. (4.34) can be rewritten as follows:

$$\begin{aligned} \sigma_{PH}^2 &= \sigma_H^2 \times trace(P' \times P) / n \\ &= \sigma_H^2 \times trace\{[(I - \boldsymbol{\lambda} \times W)^{-1}]' \times (I - \boldsymbol{\lambda} \times W)^{-1}\} / n \end{aligned} \tag{4.35}$$

If we define: $\boldsymbol{Tr} = trace\{[(I - \boldsymbol{\lambda} \times W)^{-1}]' \times (I - \boldsymbol{\lambda} \times W)^{-1}\} / n$, σ_{PH}^2 can be expressed as: $\boldsymbol{Tr} \times \sigma_H^2$.

Turning now to the conditional distribution of non-land costs C^{NL}, things become more complicated. From the relationship shown in Eq. (4.7), we can obtain the conditional distribution of C^{NL} by transforming the unconditional distribution of θ_L and P_H, since a linear combination of two normal distributions is still a normal distribution. Transforming the elements operations into matrix notation, and noting that there would be no covariance term due to the independence assumption of θ_L and P_H, we would have the mean of C^{NL}, $\boldsymbol{\mu}_C$, and the variance of C^{NL}, σ_C^2, as follows:

$$\boldsymbol{\mu}_C = diag(H') \times X_H \times \boldsymbol{\beta}_H - (H\ ./\ PR)\ .\times (\boldsymbol{\beta}_{Tobit}\, X_L + r) \tag{4.36}$$

$$\sigma_C^2 = (H' \times H / n) \times \boldsymbol{Tr} \times \sigma_H^2 + [(H\ ./\ PR)' \times (H\ ./\ PR\ )/\ n] \times \sigma_L^2 \tag{4.37}$$

where *diag(H')* denotes a diagonal matrix with elements of the vector H, " ./ " and " .× " denote the dot operations. Now, with the information of the mean and variance of θ_L, P_H and C^{NL}, we could construct the full-information log-likelihood function, denoted by $ln\boldsymbol{L}$, as follows:

$$\begin{aligned} ln\boldsymbol{L} = & [- \frac{n}{2} \times log(2 \times \boldsymbol{\pi}) - \frac{n}{2} \times log(\sigma_L^2) - \frac{1}{2 \times \sigma_L^2} \times (\theta_L - \boldsymbol{\mu}_L)' \times (\theta_L - \boldsymbol{\mu}_L)] \\ & + [- \frac{n}{2} \times log(2 \times \boldsymbol{\pi}) - \frac{n}{2} \times log(\sigma_H^2) \\ & - \frac{1}{2 \times \sigma_H^2} \times (P_H - X_H \times \boldsymbol{\beta}_H)' \times (I - \boldsymbol{\lambda} \times W)' \times (I - \boldsymbol{\lambda} \times W) \times (P_H - X_H \times \boldsymbol{\beta}_H)] \\ & + [- \frac{n}{2} \times log(2 \times \boldsymbol{\pi}) - \frac{n}{2} \times log(\sigma_C^2) - \frac{1}{2 \times \sigma_C^2} \times (C^{NL} - \boldsymbol{\mu}_C)' \times (C^{NL} - \boldsymbol{\mu}_C)] \end{aligned} \tag{4.38}$$

In Eq. (4.38), $\boldsymbol{\mu}_L$ and σ_L^2 are assumed given by the second chapter, $\boldsymbol{\beta}_H$, $\boldsymbol{\lambda}, \sigma_H^2$ are treated as unknown parameters to be estimated, and $\boldsymbol{\mu}_C, \sigma_C^2$ are given by Eqs. (4.36) and (4.37). Given the distinct spatial feature of Eq. (4.38), we call it *Spatial Full Information Maximum Likelihood* (SFIML) Estimation. Taking derivatives of $ln\boldsymbol{L}$ with respect to $\boldsymbol{\beta}_H$, $\boldsymbol{\lambda}$, and σ_H^2, we have the following first order conditions:

$$\frac{\partial \ln L}{\partial \beta_H} = \frac{1}{\sigma_H^2} \times X_H' \times (I - \boldsymbol{\lambda} \times W)' \times (I - \boldsymbol{\lambda} \times W) \times (P_H - X_H \times \boldsymbol{\beta}_H)$$

$$+ \frac{1}{\sigma_C^2} \times X_H' \times [diag(H')]' \times [\boldsymbol{\Delta} - diag(H') \times X_H \times \boldsymbol{\beta}_H] = 0 \qquad (4.39)$$

where, $\boldsymbol{\Delta} = C^{NL} + (H \,./\, PR) \,.\!\times (\boldsymbol{\beta}_{Tobit} \times X_L + u)$

$$\frac{\partial \ln L}{\partial \lambda} = \frac{1}{\sigma_H^2} \times (P_H - X_H \times \boldsymbol{\beta}_H)' \times W' \times (I - \boldsymbol{\lambda} \times W) \times (P_H - X_H \times \boldsymbol{\beta}_H)$$

$$- \frac{n}{2} \times \frac{1}{\sigma_C^2} \times \frac{\partial \sigma_C^2}{\partial \lambda} + \frac{1}{2 \times \sigma_C^4} \times \frac{\partial \sigma_C^2}{\partial \lambda} \times (C^{NL} - \boldsymbol{\mu}_C)' \times (C^{NL} - \boldsymbol{\mu}_C) = 0 \quad (4.40)$$

$$\frac{\partial \ln L}{\partial \sigma_H^2} = - \frac{n}{2} \times \frac{1}{\sigma_H^2} + \frac{1}{2 \times \sigma_H^4} \times (P_H - X_H \times \boldsymbol{\beta}_H)' \times (I - \boldsymbol{\lambda} \times W)' \times (I - \boldsymbol{\lambda} \times W)$$

$$\times (P_H - X_H \times \boldsymbol{\beta}_H) - \frac{n}{2} \times \frac{1}{\sigma_C^2} \times \frac{H' \times H}{n} \times \boldsymbol{Tr}$$

$$+ \frac{1}{2 \times \sigma_C^4} \times \frac{H' \times H}{n} \times \boldsymbol{Tr} \times (C^{NL} - \boldsymbol{\mu}_C)' \times (C^{NL} - \boldsymbol{\mu}_C) = 0 \qquad (4.41)$$

Solving Eq. (4.39), we obtain the estimator for $\hat{\beta}_H$, as follows:

$$\hat{\beta}_H = [\sigma_C^2 \times X_{HS}' \times X_{HS} + \sigma_H^2 \times X_H' \times [diag(H')]' \times diag(H') \times X_H]^{-1}$$

$$\times [\sigma_C^2 \times X_{HS}' \times P_{HS} + \sigma_H^2 \times X_H' \times [diag(H')]' \times \boldsymbol{\Delta}] \qquad (4.42)$$

where, $X_{HS} = (I - \boldsymbol{\lambda} \times W) \times X_H$, and $P_{HS} = (I - \boldsymbol{\lambda} \times W) \times P_H$. However, to solve $\boldsymbol{\lambda}$ from Eq. (4.40), we need to know $\frac{\partial \sigma_C^2}{\partial \lambda}$. By inspection, we have:

$$\frac{\partial \sigma_C^2}{\partial \lambda} = \frac{H' \times H}{n} \times \sigma_H^2 \times \frac{\partial Tr}{\partial \lambda} \tag{4.43}$$

where, $\boldsymbol{Tr} = trace\{[(I - \boldsymbol{\lambda} \times W)^{-1}]' \times (I - \boldsymbol{\lambda} \times W)^{-1}\} / n$, is defined previously.

Obviously, deriving $\frac{\partial Tr}{\partial \lambda}$ is the key to solve $\boldsymbol{\lambda}$. Totally differentiating $\boldsymbol{Tr}$ with respect to $\boldsymbol{\lambda}$, we have:

$$\begin{aligned} \mathrm{d}\boldsymbol{Tr} &= \frac{1}{n} \times \mathrm{d}\{trace\{[(I - \boldsymbol{\lambda} \times W)^{-1}]' \times (I - \boldsymbol{\lambda} \times W)^{-1}\}\} \\ &= \frac{1}{n} \times trace\{\{\mathrm{d}[(I - \boldsymbol{\lambda} \times W)^{-1}]\}' \times (I - \boldsymbol{\lambda} \times W)^{-1} \\ &\quad + [(I - \boldsymbol{\lambda} \times W)^{-1}]' \times \mathrm{d}[(I - \boldsymbol{\lambda} \times W)^{-1}]\} \qquad \text{(see 54,55)} \end{aligned} \tag{4.44}$$

Since, $\mathrm{d}[(I - \boldsymbol{\lambda} \times W)^{-1}] = - (I - \boldsymbol{\lambda} \times W)^{-1} \times \mathrm{d}(I - \boldsymbol{\lambda} \times W) \times (I - \boldsymbol{\lambda} \times W)^{-1} = (I - \boldsymbol{\lambda} \times W)^{-1} \times W \times \mathrm{d}\boldsymbol{\lambda} \times (I - \boldsymbol{\lambda} \times W)^{-1}$, we can rewrite Eq. (4.44) as:[56]

$$\begin{aligned} \mathrm{d}\boldsymbol{Tr} &= \frac{1}{n} \times trace\{[(I - \boldsymbol{\lambda} \times W)^{-1}]' \times (I - \boldsymbol{\lambda} \times W)^{-1} \times W \times \mathrm{d}\boldsymbol{\lambda} \times (I - \boldsymbol{\lambda} \times W)^{-1}\} \\ &\quad + \frac{1}{n} \times trace\{[(I - \boldsymbol{\lambda} \times W)^{-1}]' \times (I - \boldsymbol{\lambda} \times W)^{-1} \times W \times \mathrm{d}\boldsymbol{\lambda} \times (I - \boldsymbol{\lambda} \times W)^{-1}\} \\ &= \frac{2}{n} \times trace\{(I - \boldsymbol{\lambda} \times W)^{-1} \times [(I - \boldsymbol{\lambda} \times W)^{-1}]' \times (I - \boldsymbol{\lambda} \times W)^{-1} \times W \times \mathrm{d}\boldsymbol{\lambda}\} \\ &= \frac{2}{n} \times trace\{(I - \boldsymbol{\lambda} \times W)^{-1} \times [(I - \boldsymbol{\lambda} \times W)^{-1}]' \times (I - \boldsymbol{\lambda} \times W)^{-1} \times W\} \times \mathrm{d}\boldsymbol{\lambda} \end{aligned}$$

(see 57,58) (4.45)

Since $\mathrm{d}\boldsymbol{Tr} = \frac{\partial Tr}{\partial \lambda} \times \mathrm{d}\boldsymbol{\lambda}$, from Eq. (4.45), we have:

$$\frac{\partial Tr}{\partial \lambda} = \frac{2}{n} \times trace\{(I - \boldsymbol{\lambda} \times W)^{-1} \times [(I - \boldsymbol{\lambda} \times W)^{-1}]' \times (I - \boldsymbol{\lambda} \times W)^{-1} \times W\} \tag{4.46}$$

54 Here we use the property: $\mathrm{d}[trace(U)] = trace[\mathrm{d}(U)]$, and $\mathrm{d}(U') = [\mathrm{d}(U)]'$, where U is any matrix.
55 Here we use the property: $\mathrm{d}(U \times V) = \mathrm{d}(U) \times V + U \times \mathrm{d}(V)$, where U and V are any matrices with the appropriate dimension for matrix multiplication.
56 Here we use the property: $\mathrm{d}(U^{-1}) = - U^{-1} \times \mathrm{d}(U) \times U^{-1}$, where U is any invertible matrix.
57 Here we use the property: $trace(U \times V) = trace(V \times U)$.
58 Recall that $\boldsymbol{\lambda}$ is a scalar.

Now, together with Eqs. (4.43) and (4.46), we can rewrite Eq. (4.40) as follows:

$$\frac{\partial \ln L}{\partial \lambda} = \frac{1}{\sigma_H^2} \times (P_H - X_H \times \boldsymbol{\beta}_H)' \times W' \times (I - \boldsymbol{\lambda} \times W) \times (P_H - X_H \times \boldsymbol{\beta}_H)$$

$$- \frac{n}{2} \times \frac{1}{\sigma_C^2} \times \frac{H' \times H}{n} \times \sigma_H^2 \times \frac{2}{n} \times \boldsymbol{Tr_\lambda}$$

$$+ \frac{1}{2 \times \sigma_C^4} \times \frac{H' \times H}{n} \times \sigma_H^2 \times \frac{2}{n} \times \boldsymbol{Tr_\lambda} \times (C^{NL} - \boldsymbol{\mu}_C)' \times (C^{NL} - \boldsymbol{\mu}_C) = 0 \quad (4.47)$$

where, $\boldsymbol{Tr_\lambda} = trace\{(I - \boldsymbol{\lambda} \times W)^{-1} \times [(I - \boldsymbol{\lambda} \times W)^{-1}]' \times (I - \boldsymbol{\lambda} \times W)^{-1} \times W\}$. Since it is difficult to obtain an analytical expression for $\boldsymbol{\lambda}$, we need to numerically solve for the estimator of $\hat{\lambda}$.

The estimator for σ_H^2 is also difficult to derive. Substituting Eqs. (4.36) and (4.37) into Eq. (4.41) and expanding all the terms, after some rearrangement, we can rewrite Eq. (4.41) as:

$$[- 2 \times n \times (\frac{H' \times H}{n})^2 \times \boldsymbol{Tr}^2] \times \sigma_H^6 + \{- 3 \times n \times \frac{H' \times H}{n} \times \boldsymbol{Tr} \times \boldsymbol{\Omega} + (\frac{H' \times H}{n})^2 \times$$

$$\boldsymbol{Tr}^2 \times \boldsymbol{\Phi} + \frac{H' \times H}{n} \times \boldsymbol{Tr} \times [\boldsymbol{\Delta} - diag(H') \times X_H \times \boldsymbol{\beta}_H]' \times [\boldsymbol{\Delta} - diag(H') \times X_H \times \boldsymbol{\beta}_H]\} \times \sigma_H^4$$

$$+ [- n \times \boldsymbol{\Omega}^2 - 2 \times \frac{H' \times H}{n} \times \boldsymbol{Tr} \times \boldsymbol{\Omega} \times \boldsymbol{\Phi}] \times \sigma_H^2 + \boldsymbol{\Omega}^2 \times \boldsymbol{\Phi} = 0 \quad (4.48)$$

where, $\boldsymbol{\Omega} = [(H ./ PR)' \times (H ./ PR) / n] \times \sigma_L^2$, and,

$\boldsymbol{\Phi} = (P_H - X_H \times \boldsymbol{\beta}_H)' \times (I - \boldsymbol{\lambda} \times W)' \times (I - \boldsymbol{\lambda} \times W) \times (P_H - X_H \times \boldsymbol{\beta}_H)$. Now, we define:

$$a = - 2 \times n \times (\frac{H' \times H}{n})^2 \times \boldsymbol{Tr}^2 \quad (4.48a)$$

$$b = -3 \times n \times \frac{H' \times H}{n} \times \boldsymbol{Tr} \times \boldsymbol{\Omega} + (\frac{H' \times H}{n})^2 \times \boldsymbol{Tr}^2 \times \boldsymbol{\Phi} + \frac{H' \times H}{n} \times \boldsymbol{Tr}$$

$$\times [\boldsymbol{\Delta} - diag(H') \times X_H \times \boldsymbol{\beta}_H]' \times [\boldsymbol{\Delta} - diag(H') \times X_H \times \boldsymbol{\beta}_H] \quad (4.48b)$$

$$c = -n \times \boldsymbol{\Omega}^2 - 2 \times \frac{H' \times H}{n} \times \boldsymbol{Tr} \times \boldsymbol{\Omega} \times \boldsymbol{\Phi} \quad (4.48c)$$

$$d = \boldsymbol{\Omega}^2 \times \boldsymbol{\Phi} \quad (4.48d)$$

With these definitions in Eqs. (4.48a) to (4.48d), the seemingly complex Eq. (4.48) is simply a cubic polynomial of σ_H^2, which can be expressed as follows:

$$a \times (\sigma_H^2)^3 + b \times (\sigma_H^2)^2 + c \times \sigma_H^2 + d = 0 \quad (4.49)$$

The solution of Eq. (4.49) gives us the estimator for σ_H^2 as follows:[59]

$$\hat{\sigma}_H^2 = \frac{\Gamma^{1/3}}{6 \times a} - \frac{2 \times a \times c - 2 \times b^2 / 3}{a \times \Gamma^{1/3}} - \frac{b}{3 \times a} \quad (4.50)$$

where, $\boldsymbol{\Gamma} = 36 \times a \times b \times c - 108 \times a^2 \times d - 8 \times b^3 + 12 \times a \times (12 \times a \times c^3 - 3 \times b^2 \times c^2 - 54 \times a \times b \times c \times d + 81 \times a^2 \times d^2 + 12 \times b^3 \times d)^{1/2}$

Once we have obtained all the estimators of the SFIML estimation, we need to estimate the precision of our results. Hence we need to derive the variance - covariance matrix of the SFIML function. Denoting all the estimated parameters in the SFIML as θ_0, by the property of maximum likelihood estimators, we know that, asymptotically, the variance - covariance matrix would be $\{I(\theta_0)\}^{-1}$, where $I(\theta_0)$ is the information matrix that is defined as follows:

59 Theoretically, there would be three roots for a standard cubic polynomial: one real root and two imaginary roots. In this study, we only consider the real root.

$$I(\theta_0) = -\,\boldsymbol{E}_0[\frac{\partial^2 \ln L}{\partial\theta_0 \partial\theta_0'}] \qquad (4.51)$$

where $\boldsymbol{E}_0$ stands for the expectation of the hessian matrix evaluated at θ_0. $I(\theta_0)$, therefore, is presented as follows (see Appendix for its proof):

$$I(\theta_0) = \begin{bmatrix} I_{11} & 0 & 0 \\ 0 & I_{22} & I_{23} \\ 0 & I_{32} & I_{33} \end{bmatrix} \qquad (4.52)$$

where:

$$I_{11} = \frac{1}{\sigma_H^2} \times X_H' \times (I - \boldsymbol{\lambda} \times W)' \times (I - \boldsymbol{\lambda} \times W) \times X_H'$$

$$+ \frac{1}{\sigma_C^2} \times X_H' \times [diag(H')]' \times [diag(H')] \times X_H \qquad (4.52a)$$

$$I_{22} = trace\{[(I - \boldsymbol{\lambda} \times W)^{-1}]' \times W' \times W \times (I - \boldsymbol{\lambda} \times W)^{-1}\}$$

$$+ \frac{2 \times \sigma_H^4}{n \times \sigma_C^4} \times (\frac{H' \times H}{n})^2 \times (\boldsymbol{Tr_\lambda})^2 \qquad (4.52b)$$

$$I_{23} = I_{32} = \frac{1}{\sigma_H^2} \times trace\{[(I - \boldsymbol{\lambda} \times W)^{-1}]' \times W'\} + \frac{\sigma_H^2 \times \frac{H' \times H}{n} \times Tr}{\sigma_C^4} \times \frac{H' \times H}{n} \times \boldsymbol{Tr_\lambda}$$

$$(4.52c)$$

$$I_{33} = \frac{n}{2 \times \sigma_H^4} + \frac{n}{2 \times \sigma_C^4} \times (\frac{H' \times H}{n})^2 \times \boldsymbol{Tr}^2 \qquad (4.52d)$$

Inverting $I(\theta_0)$ in Eq. (4.52), we obtain the variance - covariance matrix. With this variance - covariance matrix, along with the estimators shown in Eqs. (4.42), (4.47) and (4.50), we can estimate our SFIML model in an iterative manner.[60] The

60 On a workstation with 3.0 Ghz quad core CPU and 8 GB RAM, it takes

estimation results are shown in Tables 4.8, 4.9, and 4.10, for the cases using deflation only by HPI, monthly time dummy, and both deflation and dummy, respectively. Overall, our SFIML estimation works fairly well. The sign of the three key Project-Attribute Variables are all consistent with our expectation and they are all statistically significant. Only one variable in the housing unit attributes, "Close to street (1=yes)," has a positive sign while we would expect it to be negative. It may be due to the 4.project data set that has been used.[61] For the cases using deflation only by HPI, monthly time dummy, and a mixed use of HPI deflation and monthly dummy, it appears that the mixed one performs the best in the sense that it has the largest log-likelihood value and smallest sum of square of the estimation residuals. However, the estimation coefficients and the corresponding *t*-values are roughly the same in all cases. With these estimation results, we could implement our hypothesis test on the validity of the linkage between the land and housing markets, as well as the profit-maximization behavior of the property developer. We use the estimation results of the standard SEM in the separate housing market as a constraint on our SFIML estimation. Since we have derived the variance - covariance matrix, a natural candidate for the test would be the Lagrange multiplier (LM) test.

---

about 8 to 24 hours to run one round of estimation with the step size of 0.0001 for the three different scenarios. It would be much faster with a larger step size such as 0.01.

61 If we redo the separate estimation on the housing market using standard SEM with the same 4-project housing data set rather than the original 6-project one, we get the same sign in this variable.

Table 4.8
SFIML Estimation of Unit Housing Retail Sales Price, P_H (￥/ m^2), Using HPI Deflation

Variable	Coefficient	t-statistic	p-value
Intercept	18426.264	2.067	0.037
Project-Attribute Variables ("PR" and "X_1")			
Plot Ratio	-1378.091	-1.718	0.086
ln(Subway station proximity)	-1359.144	-1.769	0.077
Road aggregate	293.246	1.006	0.314
Housing-Unit-Attribute Variables ("H" and "X_3")			
Floor	2.576	2.390	0.017
Inner view (1=yes)	299.652	15.291	0.000
Outer view (1=yes)	92.645	4.697	0.000
Close to street (1=yes)	45.295	2.170	0.030
North (1=yes)	-77.571	-3.252	0.001
North East (1=yes)	-43.734	-1.262	0.207
South East (1=yes)	-2.7109	-0.142	0.887
South (1=yes)	33.223	1.033	0.302
South West (1=yes)	-18.392	-0.666	0.506
West (1=yes)	205.657	5.758	0.000
Distance to inner source	-0.665	-2.688	0.007
Housing unit area	0.654	1.727	0.084
Pay in cash (1=yes)	-73.831	-5.685	0.000
λ	0.990	203.235	0.000
Adjusted R-square		0.824	
sigma^2		21061.913	
log-likelihood		-26820.193	

Dependent variable: Deflated per unit housing retail sales price

We set H_0 as "estimation parameters of the separate SEM models in housing market are valid constraints on the estimation parameters of SFIML." Denoting $\hat{\beta}_{HR}$ as the constraint, we can express the LM statistic as follows:

Table 4.9
SFIML Estimation of Unit Housing Retail Sales Price, P_H (¥/ m^2), Using Monthly Time Dummy

Variable	Coefficient	t-statistic	p-value
Intercept	26730.837	2.164	0.030
Project-Attribute Variables (*"PR" and "X_1"*)			
Plot Ratio	-1857.599	-1.671	0.095
ln(Subway station proximity)	-2021.866	0.815	0.415
Road aggregate	329.139	-1.900	0.057
Housing-Unit-Attribute Variables (*"H" and "X_3"*)			
Floor	4.583	3.069	0.002
Inner view (1=yes)	399.713	14.721	0.000
Outer view (1=yes)	122.882	4.496	0.000
Close to street (1=yes)	67.815	2.340	0.019
North (1=yes)	-102.711	-3.101	0.002
North East (1=yes)	-85.622	-1.778	0.075
South East (1=yes)	-17.712	-0.668	0.504
South (1=yes)	41.176	0.924	0.356
South West (1=yes)	-38.296	-1.000	0.317
West (1=yes)	265.153	5.358	0.000
Distance to inner source	-0.909	-2.651	0.008
Housing unit area	0.845	1.600	0.110
Pay in cash (1=yes)	-89.862	-4.993	0.000
Time trend	1.951	0.673	0.501
λ	0.990	203.235	0.000
Adjusted R-square		0.805	
sigma^2		40426.258	
log-likelihood		-27546.491	

Dependent variable: Per unit housing retail sales price

$$LM = \{\frac{1}{\sigma_H^2} \times X_H' \times (I - \lambda \times W)' \times (I - \lambda \times W) \times (P_H - X_H \times \hat{\beta}_{HR}) + \frac{1}{\sigma_C^2} \times X_H' \times$$

$$[diag(H')]' \times [\Delta - diag(H') \times X_H \times \hat{\beta}_{HR}\}' \times \{I(\hat{\beta}_{HR})\}^{-1} \times \{\frac{1}{\sigma_H^2} \times X_H' \times (I - \lambda \times W)' \times (I -$$

Table 4.10
SFIML Estimation of Unit Housing Retail Sales Price, P_H (￥/ m^2), Using Both HPI Deflation and Monthly Dummy

Variable	Coefficient	t-statistic	p-value
Intercept	18265.178	2.049	0.040
Project-Attribute Variables ("PR" and "X_1")			
Plot Ratio	-1396.282	-1.741	0.082
ln(Subway station proximity)	-1351.383	-1.760	0.078
Road aggregate	301.562	1.035	0.301
Housing-Unit-Attribute Variables ("H" and "X_3")			
Floor	2.595	2.408	0.016
Inner view (1=yes)	299.412	15.281	0.000
Outer view (1=yes)	92.455	4.688	0.000
Close to street (1=yes)	46.378	2.218	0.027
North (1=yes)	-76.152	-3.186	0.001
North East (1=yes)	-45.872	-1.320	0.187
South East (1=yes)	-3.249	-0.170	0.865
South (1=yes)	33.757	1.050	0.294
South West (1=yes)	-19.351	-0.700	0.484
West (1=yes)	205.136	5.744	0.000
Distance to inner source	-0.669	-2.703	0.007
Housing unit area	0.616	1.618	0.106
Pay in cash (1=yes)	-73.522	-5.661	0.000
Time trend	1.835	0.877	0.380
λ	0.990	203.235	0.000
Adjusted R-square		0.829	
sigma^2		21052.352	
log-likelihood		-26813.661	

Dependent variable: Deflated per unit housing retail sales price

$$\lambda \times W) \times (P_H - X_H \times \hat{\beta}_{HR}) + \frac{1}{\sigma_C^2} \times X_H' \times [diag(H')]' \times [\Delta - diag(H') \times X_H \times \hat{\beta}_{HR}\} \quad (4.53)$$

As a result, the values of the LM statistics are 0.411, 0.338, and 0.262 for the cases using HPI deflation, monthly time dummy, and both HPI deflation and monthly dummy, respectively. Therefore, it fails to reject the null hypothesis at the 5% significance level for all cases.

We consider the "fail to reject" results in the LM test as a positive signal in this study. For the separate estimation in the housing market using standard SEM approach, we have imposed no information about the land market, nor the profit maximization behavior of the property developer. However, the property development activity is a complete process in the sense that it incorporates the factor market (i.e., the land market), the commodity market (i.e., the housing market), as well as the profit maximization behavior all together. There has to be a way, at least in the theoretical level, to link them together with respect to the various characteristics that could be capitalized in the housing price. Our SFIML approach, as shown above, has successfully linked these three components together, from theory to empirical practice. Essentially, the "fail to reject" results in the hypothesis test (i.e., the LM test) have statistically proved our theoretical assumption using real world data; the results have also demonstrated the validity and robustness of the hedonic theory with both input and output. In addition, since the property developer acquires raw land parcels via English auction, our approach has also shown that the derived true valuation of land works fairly well with the SFIML estimation, which has confirmed and justified the use of the derived true valuation in the hedonic price estimation.

6. Concluding Remarks

In this study, we have demonstrated two methods to implement the hypothesis test on the hedonic price estimation with both input and output, which has been applied to a Chinese regional land market and housing market. We focus on the profit maximization behavior of the property developer, which is the key role to link the factor market (i.e., the land market) and the commodity market (i.e., the housing market), as well as the profit maximization behavior in the property development activity. Although the housing market is assumed to be competitive and no economic profit exists, a positive profit is allowed from the premium due to the difference between the buyer's true valuation and the actual sales price in the English auction where the raw land parcel is traded with the local government. With the developer's true valuation of land derived in Chapter 1, we have calculated a non-land costs in the property development process aside from the land cost.

Two methods are employed to conduct the hypothesis test. A set of partial derivatives of the profit function with respect to various characteristics give us the relationship between the marginal valuations in the land and housing markets, which then present us the link between the estimation parameters in these two markets, and also play the role as constraints on the estimation parameters. We also use a joint estimation approach which considers the land market, the housing market and the property developer's profit maximization behavior all together in the estimation. We then use the results in the corresponding separate estimation in the housing market as constraints on the parameters. In the separate estimation model, our results reject the null hypothesis that the calculated constraints are valid, but it is highly possible to be due to the fact that our limited number of housing projects in the data set might be a poor representative of all the land parcels, since the hedonic parameters may vary from neighborhood to neighborhood as Ellickson (1981) has pointed out. In the joint estimation model, our results fail to reject the null hypothesis, which we consider to be a positive signal to confirm and justify the theoretical linkage (i.e., our linked markets assumption) in the hedonic price estimation.

References

Al Refai, A. S. A. M., 1994. Analysis of tax incidence within a simultaneous system of hedonic demand equations: a case study of the Los Angeles housing market. Ph.D dissertation, The Claremont Graduate School.

Chow, G. C., 1968. Two methods of computing full-information maximum likelihood estimates in simultaneous stochastic equations. International Economic Review 9 (1) (February), 100–112.

Ellickson, B., 1981. An alternative test of the hedonic theory of housing markets. Journal of Urban Economics 9 (1) (January), 56–79.

Eisenpress, H., 1962. Note on the computation of full-information maximum-likelihood estimates of coefficients of a simultaneous system. Econometrica 30 (2) (April), 343–348.

Heal, G. M., 2001. Bundling public and private goods. Columbia Business School; National Bureau of Economic Research (NBER), Working Paper Series. Available at: SSRN: http://ssrn.com/abstract=279097 or doi: 10.2139/ssrn.279097.

Palmquist, R. B., 1989. Land as a differentiated factor of production: a hedonic model and its implications for welfare measurement. Land Economics 65 (February), 23–28.

Rosen, S., 1974. Hedonic prices and implicit markets: product differentiation in only competition. Journal of Political Economy 82 (1), 34–55.

Santerre, R. E., 1985. Spatial differences in the demands for local public goods. Land Economics 61 (2) (May), 119–128.

Wu, J., 2006. Environmental amenities, urban sprawl, and community characteristics. Journal of Environmental Economics and Management 52 (2) (September), 527–547, doi: 10.1016/j.jeem.2006.03.003.

APPENDIX

Derivation of the Information Matrix (4.52)

Now, let's look at the derivation of the information matrix, $I(\theta_0)$, in the SFIML. In order to obtain the information matrix, the first step is to derive the hessian matrix of $ln\boldsymbol{L}$. While some of the second order derivatives are straight forward to obtain, some are a bit more complex to derive. The second order condition for $\boldsymbol{\beta}_H$ is relatively straight forward to derive. Differentiating Eq. (4.40) with respect to $\boldsymbol{\beta}_H$, we have:

$$\frac{\partial^2 \ln L}{\partial {\beta_H}^2} = - \frac{1}{\sigma_H^2} \times X_H' \times (I - \boldsymbol{\lambda} \times W)' \times (I - \boldsymbol{\lambda} \times W) \times X_H$$

$$- \frac{1}{\sigma_C^2} \times X_H' \times [diag(H')]' \times [diag(H')] \times X_H \qquad \text{(A.1)}$$

The second order condition for $\boldsymbol{\lambda}$ is difficult to obtain. Noting that both σ_C^2 and $\boldsymbol{Tr}_\lambda$ are functions of $\boldsymbol{\lambda}$, by chain rule, we have:

$$\frac{\partial^2 \ln L}{\partial \lambda^2} = - \frac{1}{\sigma_H^2} \times (P_H - X_H \times \boldsymbol{\beta}_H)' \times W' \times W \times (P_H - X_H \times \boldsymbol{\beta}_H)$$

$$+ \left(\frac{\sigma_H^2}{\sigma_C^4} \times \frac{\partial \sigma_C^2}{\partial \lambda} \times \frac{H' \times H}{n} \times \boldsymbol{Tr}_\lambda - \frac{\sigma_H^2}{\sigma_C^2} \times \frac{H' \times H}{n} \times \frac{\partial Tr_\lambda}{\partial \lambda}\right)$$

$$- \frac{2 \times \sigma_H^2}{\sigma_C^6} \times \frac{\partial \sigma_C^2}{\partial \lambda} \times \frac{H' \times H}{n} \times \frac{1}{n} \times \boldsymbol{Tr}_\lambda \times (C^{NL} - \boldsymbol{\mu}_C)' \times (C^{NL} - \boldsymbol{\mu}_C)$$

$$+ \frac{\sigma_H^2}{\sigma_C^4} \times \frac{H' \times H}{n} \times \frac{1}{n} \times \frac{\partial Tr_\lambda}{\partial \lambda} \times (C^{NL} - \boldsymbol{\mu}_C)' \times (C^{NL} - \boldsymbol{\mu}_C) \qquad \text{(A.2)}$$

where, $\frac{\partial \sigma_C^2}{\partial \lambda} = \frac{H' \times H}{n} \times \sigma_H^2 \times \frac{2}{n} \times \boldsymbol{Tr}_\lambda$, as shown in Eqs. (4.43) and (4.46). To obtain $\frac{\partial Tr_\lambda}{\partial \lambda}$, denoting $\zeta = I - \boldsymbol{\lambda} \times W$, and totally differentiating $\boldsymbol{Tr}_\lambda$ with respect to $\boldsymbol{\lambda}$, we have:

$$d\{trace[\zeta^{-1} \times (\zeta^{-1})' \times \zeta^{-1} \times W]\} = trace\{d[\zeta^{-1} \times (\zeta^{-1})' \times \zeta^{-1} \times W]\}$$

$$= trace[d(\zeta^{-1}) \times (\zeta^{-1})' \times \zeta^{-1} \times W] + trace\{\zeta^{-1} \times [d(\zeta^{-1})]' \times \zeta^{-1} \times W\}$$

$$+ trace[\zeta^{-1} \times (\zeta^{-1})' \times d(\zeta^{-1}) \times W]$$

$$= trace[- \zeta^{-1} \times d(\zeta) \times \zeta^{-1} \times (\zeta^{-1})' \times \zeta^{-1} \times W]$$

$$+ trace\{- \zeta^{-1} \times (\zeta^{-1})' \times [d(\zeta)]' \times (\zeta^{-1})' \times \zeta^{-1} \times W\}$$

$$+ trace[- \zeta^{-1} \times (\zeta^{-1})' \times \zeta^{-1} \times d(\zeta) \times \zeta^{-1} \times W] \tag{A.3}$$

Noting that $d(\zeta) = d(I - \boldsymbol{\lambda} \times W) = - W \times d\boldsymbol{\lambda}$, and using properties "$trace(A \times B) = trace(B \times A)$" and "$trace(A) = trace(A')$," after rearrangement, Eq. (A.3) can be rewritten as:

$$d\boldsymbol{Tr}_{\lambda} = trace[\zeta^{-1} \times (\zeta^{-1})' \times \zeta^{-1} \times W \times \zeta^{-1} \times W] \times d\boldsymbol{\lambda}$$

$$+ trace[\zeta^{-1} \times (\zeta^{-1})' \times W' \times (\zeta^{-1})' \times \zeta^{-1} \times W] \times d\boldsymbol{\lambda}$$

$$+ trace[\zeta^{-1} \times W \times \zeta^{-1} \times (\zeta^{-1})' \times \zeta^{-1} \times W] \times d\boldsymbol{\lambda} \tag{A.4}$$

Therefore, we have:

$$\frac{\partial Tr_{\lambda}}{\partial \lambda} = trace[\zeta^{-1} \times (\zeta^{-1})' \times \zeta^{-1} \times W \times \zeta^{-1} \times W]$$

$$+ trace[\zeta^{-1} \times (\zeta^{-1})' \times W' \times (\zeta^{-1})' \times \zeta^{-1} \times W]$$

$$+ trace[\zeta^{-1} \times W \times \zeta^{-1} \times (\zeta^{-1})' \times \zeta^{-1} \times W] \tag{A.5}$$

Substituting $\frac{\partial \sigma_C^2}{\partial \lambda}$ and Eq. (A.5) into Eq. (A.2), we have:

$$\frac{\partial^2 \ln L}{\partial \lambda^2} = - \frac{1}{\sigma_H^2} \times (P_H - X_H \times \boldsymbol{\beta}_H)' \times W' \times W \times (P_H - X_H \times \boldsymbol{\beta}_H)$$

$$+ \left(\frac{\sigma_H^2}{\sigma_C^4} \times \frac{H' \times H}{n} \times \sigma_H^2 \times \frac{2}{n} \times \boldsymbol{Tr_\lambda} \times \frac{H' \times H}{n} \times \boldsymbol{Tr_\lambda} - \frac{\sigma_H^2}{\sigma_C^2} \times \frac{H' \times H}{n} \times \boldsymbol{Tr_{\lambda\lambda}}\right)$$

$$- \frac{2 \times \sigma_H^2}{\sigma_C^6} \times \frac{H' \times H}{n} \times \sigma_H^2 \times \frac{2}{n} \times \boldsymbol{Tr_\lambda} \times \frac{H' \times H}{n} \times \frac{1}{n} \times \boldsymbol{Tr_\lambda} \times (C^{NL} - \boldsymbol{\mu}_C)' \times (C^{NL} - \boldsymbol{\mu}_C)$$

$$+ \frac{\sigma_H^2}{\sigma_C^4} \times \frac{H' \times H}{n} \times \frac{1}{n} \times \boldsymbol{Tr_{\lambda\lambda}} \times (C^{NL} - \boldsymbol{\mu}_C)' \times (C^{NL} - \boldsymbol{\mu}_C) \qquad \text{(A.6)}$$

where, $\boldsymbol{Tr_{\lambda\lambda}} = trace[\zeta^{-1} \times (\zeta^{-1})' \times \zeta^{-1} \times W \times \zeta^{-1} \times W] + trace[\zeta^{-1} \times (\zeta^{-1})' \times W' \times (\zeta^{-1})' \times \zeta^{-1} \times W] + trace[\zeta^{-1} \times W \times \zeta^{-1} \times (\zeta^{-1})' \times \zeta^{-1} \times W]$, and $\zeta = I - \boldsymbol{\lambda} \times W$, as shown in Eq. (A.5).

The second order condition for σ_H^2 is shown as follows:

$$\frac{\partial^2 \ln L}{\partial(\sigma_H^2)^2} = \frac{n}{2 \times \sigma_H^4} - \frac{1}{\sigma_H^6} \times (P_H - X_H \times \boldsymbol{\beta}_H)' \times (I - \boldsymbol{\lambda} \times W)' \times (I - \boldsymbol{\lambda} \times W)$$

$$\times (P_H - X_H \times \boldsymbol{\beta}_H) + \frac{n}{2 \times \sigma_C^4} \times \left(\frac{H' \times H}{n}\right)^2 \times \boldsymbol{Tr}^2$$

$$- \frac{1}{\sigma_C^6} \times \left(\frac{H' \times H}{n}\right)^2 \times \boldsymbol{Tr}^2 \times (C^{NL} - \boldsymbol{\mu}_C)' \times (C^{NL} - \boldsymbol{\mu}_C) \qquad \text{(A.7)}$$

The cross-second-order conditions of $\frac{\partial^2 \ln L}{\partial\beta\partial\sigma_H^2}$ (= $\frac{\partial^2 \ln L}{\partial\sigma_H^2\partial\beta}$) and $\frac{\partial^2 \ln L}{\partial\beta_H\partial\lambda}$ (= $\frac{\partial^2 \ln L}{\partial\lambda\partial\beta_H}$) are shown in Eqs. (A.8) and (A.9), respectively, as follows:

$$\frac{\partial^2 \ln L}{\partial\beta\partial\sigma_H^2} = \frac{\partial^2 \ln L}{\partial\sigma_H^2\partial\beta} = -\frac{1}{\sigma_H^4} \times X_H{}' \times (I - \boldsymbol{\lambda} \times W)' \times (I - \boldsymbol{\lambda} \times W) \times (P_H - X_H \times \boldsymbol{\beta}_H)$$

$$- \frac{1}{\sigma_C^4} \times \frac{H' \times H}{n} \times \boldsymbol{Tr} \times X_H' \times [diag(H')]' \times [\boldsymbol{\Delta} - diag(H') \times X_H \times \boldsymbol{\beta}_H]$$

$$\text{(A.8)}$$

$$\frac{\partial^2 \ln L}{\partial\beta_H \partial\lambda} = \frac{\partial^2 \ln L}{\partial\lambda \partial\beta_H} = -\frac{2}{\sigma_H^2} \times X_H' \times W' \times (I - \boldsymbol{\lambda} \times W) \times (P_H - X_H \times \boldsymbol{\beta}_H)$$

$$-\frac{2}{\sigma_C^4} \times \frac{H' \times H}{n} \times \sigma_H^2 \times \frac{1}{n} \times \boldsymbol{Tr_\lambda} \times X_H' \times [diag(H')]' \times (C^{NL} - \boldsymbol{\mu}_C) \qquad \text{(A.9)}$$

The cross-second-order condition of $\frac{\partial^2 \ln L}{\partial\lambda \partial\sigma_H^2}$ (= $\frac{\partial^2 \ln L}{\partial\sigma_H^2 \partial\lambda}$) is a bit more complicated. Noting that "$d(\frac{U}{V}) = \frac{dU \times V + U \times dV}{V^2}$," along with the chain rule, we have:

$$\frac{\partial^2 \ln L}{\partial\lambda \partial\sigma_H^2} = \frac{\partial^2 \ln L}{\partial\sigma_H^2 \partial\lambda}$$

$$= -\frac{1}{\sigma_H^4} \times (P_H - X_H \times \boldsymbol{\beta}_H)' \times W' \times (I - \boldsymbol{\lambda} \times W) \times (P_H - X_H \times \boldsymbol{\beta}_H)$$

$$-\frac{\sigma_C^2 - \sigma_H^2 \times \frac{H' \times H}{n} \times Tr}{\sigma_C^4} \times \frac{H' \times H}{n} \times \boldsymbol{Tr_\lambda}$$

$$+\frac{\sigma_C^2 - 2 \times \sigma_H^2 \times \frac{H' \times H}{n} \times Tr}{\sigma_C^6} \times \frac{H' \times H}{n} \times \frac{1}{n} \times \boldsymbol{Tr_\lambda} \times (C^{NL} - \boldsymbol{\mu}_C)' \times (C^{NL} - \boldsymbol{\mu}_C)$$

(A.10)

Now, we take expectation on Eqs. (A.1), (A.6), (A.7) to (A.9), and (A.10). Every term in Eq. (A.1) is non-stochastic, hence its expectation remains the same. Thus,

$$E_0[\frac{\partial^2 \ln L}{\partial{\beta_H}^2}] = -\frac{1}{\sigma_H^2} \times X_H' \times (I - \boldsymbol{\lambda} \times W)' \times (I - \boldsymbol{\lambda} \times W) \times X_H'$$

$$-\frac{1}{\sigma_C^2} \times X_H' \times [diag(H')]' \times X_H' \times [diag(H')]' \qquad \text{(A.11)}$$

To obtain $E_0[\frac{\partial^2 \ln L}{\partial\lambda^2}]$, we need to make expectation on each of the four terms in Eq.

(A.6) one by one. For the first term which we denote by $\left.\frac{\partial^2 \ln L}{\partial \lambda^2}\right|_1$, "- $\frac{1}{\sigma_H^2}$ × $(P_H - X_H \times \boldsymbol{\beta}_H)' \times W' \times W \times (P_H - X_H \times \boldsymbol{\beta}_H)$," we need to substitute $P_H - X_H \times \boldsymbol{\beta}_H = (I - \boldsymbol{\lambda} \times W)^{-1} \times e_H$ from Eq. (4.33) into it, as follows:

$$\left.\frac{\partial^2 \ln L}{\partial \lambda^2}\right|_1 = -\frac{1}{\sigma_H^2} \times e_H' \times [(I - \boldsymbol{\lambda} \times W)^{-1}]' \times W' \times W \times (I - \boldsymbol{\lambda} \times W)^{-1} \times e_H \quad \text{(A.12)}$$

Denoting $M = W \times (I - \boldsymbol{\lambda} \times W)^{-1}$, by inspection, the element in the i^{th} row of the vector $M \times e_H$ is $\sum_j^n M_{ij} \times e_{Hj}$, where j denotes the j^{th} element in the i^{th} row of the matrix M.

Therefore, the matrix multiplication $(M \times e_H)' \times (M \times e_H)$ can be expressed as $\sum_i^n (\sum_j^n M_{ij} \times e_{Hj})^2$. Taking expectation on $(M \times e_H)' \times (M \times e_H)$, we have:

$$E_0[\sum_i^n (\sum_j^n M_{ij} \times e_{Hj})^2] = \sum_i^n E_0[(\sum_j^n M_{ij} \times e_{Hj})^2] \quad \text{(A.13)}$$

Noting the fact that $E_0[\sum_j^n M_{ij} \times e_{Hj}] = 0$, using the definition of variance, Eq. (A.13) can be rewritten as follows:

$$\sum_i^n E_0[(\sum_j^n M_{ij} \times e_{Hj})^2] = \sum_i^n Var_0[\sum_j^n M_{ij} \times e_{Hj}] = \sum_i^n (\sum_j^n M_{ij}^2 \times \sigma_H^2)$$

$$= \sigma_H^2 \times \sum_i^n \sum_j^n M_{ij}^2 = \sigma_H^2 \times trace(M' \times M) \quad \text{(A.14)}$$

Therefore, combining Eqs. (A.12) and (A.14), we have:

$$E_0[\left.\frac{\partial^2 \ln L}{\partial \lambda^2}\right|_1] = -trace\{[(I - \boldsymbol{\lambda} \times W)^{-1}]' \times W' \times W \times (I - \boldsymbol{\lambda} \times W)^{-1}\} \quad \text{(A.15)}$$

The second term of Eq. (A.6), $\left.\frac{\partial^2 \ln L}{\partial \lambda^2}\right|_2$, is non-stochastic, hence, its expectation remains the same, i.e.,

$$E_0\left[\left.\frac{\partial^2 \ln L}{\partial \lambda^2}\right|_2\right] = \frac{2\times\sigma_H^4}{n\times\sigma_C^4} \times \left(\frac{H'\times H}{n}\right)^2 \times (\boldsymbol{Tr_\lambda})^2 - \frac{\sigma_H^2}{\sigma_C^2} \times \frac{H'\times H}{n} \times \boldsymbol{Tr_{\lambda\lambda}} \tag{A.16}$$

Realizing the fact that,

$$E_0[(C^{NL} - \boldsymbol{\mu}_C)' \times (C^{NL} - \boldsymbol{\mu}_C)] = n\times\sigma_C^2 \tag{A.17}$$

The expectation of the third and fourth terms of Eq. (A.6), i.e., $\left.\frac{\partial^2 \ln L}{\partial \lambda^2}\right|_3$ and $\left.\frac{\partial^2 \ln L}{\partial \lambda^2}\right|_4$, can be expressed in Eqs. (A.18) and (A.19), respectively, as follows:

$$E_0\left[\left.\frac{\partial^2 \ln L}{\partial \lambda^2}\right|_3\right] = - \frac{4\times\sigma_H^4}{n\times\sigma_C^4} \times \left(\frac{H'\times H}{n}\right)^2 \times (\boldsymbol{Tr_\lambda})^2 \tag{A.18}$$

$$E_0\left[\left.\frac{\partial^2 \ln L}{\partial \lambda^2}\right|_4\right] = \frac{\sigma_H^2}{\sigma_C^2} \times \frac{H'\times H}{n} \times \boldsymbol{Tr_{\lambda\lambda}} \tag{A.19}$$

Adding up Eqs. (A.15), (A.16), (A.18), and (A.19), we have:

$$\begin{aligned}
E_0\left[\frac{\partial^2 \ln L}{\partial \lambda^2}\right] &= - trace\{[(I - \boldsymbol{\lambda} \times W)^{-1}]' \times W' \times W \times (I - \boldsymbol{\lambda} \times W)^{-1}\} \\
&\quad + \left(\frac{2\times\sigma_H^4}{n\times\sigma_C^4} \times \left(\frac{H'\times H}{n}\right)^2 \times (\boldsymbol{Tr_\lambda})^2 - \frac{\sigma_H^2}{\sigma_C^2} \times \frac{H'\times H}{n} \times \boldsymbol{Tr_{\lambda\lambda}}\right) \\
&\quad - \frac{4\times\sigma_H^4}{n\times\sigma_C^4} \times \left(\frac{H'\times H}{n}\right)^2 \times (\boldsymbol{Tr_\lambda})^2 + \frac{\sigma_H^2}{\sigma_C^2} \times \frac{H'\times H}{n} \times \boldsymbol{Tr_{\lambda\lambda}} \\
&= - trace\{[(I - \boldsymbol{\lambda} \times W)^{-1}]' \times W' \times W \times (I - \boldsymbol{\lambda} \times W)^{-1}\} \\
&\quad - \frac{2\times\sigma_H^4}{n\times\sigma_C^4} \times \left(\frac{H'\times H}{n}\right)^2 \times (\boldsymbol{Tr_\lambda})^2
\end{aligned} \tag{A.20}$$

Realizing the fact that the term $E_0[e_H' \times e_H]$ is nothing but $n\times\sigma_H^2$, we have:

$$E_0[(P_H - X_H \times \boldsymbol{\beta}_H)' \times (I - \boldsymbol{\lambda} \times W)' \times (I - \boldsymbol{\lambda} \times W) \times (P_H - X_H \times \boldsymbol{\beta}_H)] = n \times \sigma_H^2 \tag{A.21}$$

Therefore, considering Eqs. (A.17) and (A.21), when taking expectation, Eq. (A.7) can be transformed as:

$$E_0[\frac{\partial^2 \ln L}{\partial(\sigma_H^2)^2}] = -\frac{n}{2 \times \sigma_H^4} - \frac{n}{2 \times \sigma_C^4} \times (\frac{H' \times H}{n})^2 \times \boldsymbol{Tr}^2 \tag{A.22}$$

Noting the fact that $E_0[P_H - X_H \times \boldsymbol{\beta}_H] = (I - \boldsymbol{\lambda} \times W)^{-1} \times E_0[e_H] = 0$, and also $E_0[C^{NL} - \boldsymbol{\mu}_C] = 0$, from Eq. (A.8) we have:

$$E_0[\frac{\partial^2 \ln L}{\partial \beta \partial \sigma_H^2}] = E_0[\frac{\partial^2 \ln L}{\partial \sigma_H^2 \partial \beta}] = 0 \tag{A.23}$$

In a similar manner, from Eq. (A.9), we have:

$$E_0[\frac{\partial^2 \ln L}{\partial \beta_H \partial \lambda}] = E_0[\frac{\partial^2 \ln L}{\partial \lambda \partial \beta_H}] = 0 \tag{A.24}$$

The expectation of $\frac{\partial^2 \ln L}{\partial \lambda \partial \sigma_H^2}$ (= $\frac{\partial^2 \ln L}{\partial \sigma_H^2 \partial \lambda}$) is relatively more complicated than the previous two. Again, by Eq. (4.33), we can rewrite the first term of Eq. (A.10), which we denote by $\left.\frac{\partial^2 \ln L}{\partial \lambda \partial \sigma_H^2}\right|_1$, as follows:

$$\left.\frac{\partial^2 \ln L}{\partial \lambda \partial \sigma_H^2}\right|_1 = -\frac{1}{\sigma_H^4} \times e_H' \times [(I - \boldsymbol{\lambda} \times W)^{-1}]' \times W' \times (I - \boldsymbol{\lambda} \times W) \times (I - \boldsymbol{\lambda} \times W)^{-1} \times e_H$$

$$= -\frac{1}{\sigma_H^4} \times e_H' \times [(I - \boldsymbol{\lambda} \times W)^{-1}]' \times W' \times e_H \tag{A.25}$$

Denoting $Q = W'$, along with the notation of $P = (I - \lambda \times W)^{-1}$, by inspection, the element in the i^{th} row of the vector $P \times e_H$ is $\sum_j^n P_{ij} \times e_{Hj}$, and the element in the i^{th} row of the vector $Q \times e_H$ is $\sum_j^n Q_{ij} \times e_{Hj}$, where j denotes the j^{th} element in the i^{th} row of the matrix P and Q respectively. Therefore, we can express the matrix multiplication $(P \times e_H)' \times (Q \times e_H)$ as:

$$(P \times e_H)' \times (Q \times e_H) = \sum_i^n [(\sum_j^n P_{ij} \times e_{Hj}) \times (\sum_j^n Q_{ij} \times e_{Hj})] \qquad \text{(A.26)}$$

In order to take expectation on Eq. (A.26), we note the fact that $E_0[\sum_j^n P_{ij} \times e_{Hj}] = 0$, and also $E_0[\sum_j^n Q_{ij} \times e_{Hj}] = 0$. Then, with the definition of covariance, we have:

$$E_0\{\sum_i^n [(\sum_j^n P_{ij} \times e_{Hj}) \times (\sum_j^n Q_{ij} \times e_{Hj})]\} = \sum_i^n E_0[(\sum_j^n P_{ij} \times e_{Hj}) \times (\sum_j^n Q_{ij} \times e_{Hj})]$$

$$= \sum_i^n Cov_0(\sum_j^n P_{ij} \times e_{Hj}, \sum_j^n Q_{ij} \times e_{Hj}) = \sum_i^n \{\sum_j^n \sum_j^n [Cov_0(P_{ij} \times e_{Hj}, Q_{ij} \times e_{Hj})]\}$$

$$= \sum_i^n \{\sum_j^n \sum_j^n [P_{ij} \times Q_{ij} \times Cov_0(e_{Hj}, e_{Hj})]\} = \sum_i^n [\sum_j^n \sum_j^n (P_{ij} \times Q_{ij} \times \sigma_H^2)]$$

$$= \sigma_H^2 \times \sum_i^n \sum_j^n (P_{ij} \times Q_{ij})$$

$$= \sigma_H^2 \times trace(P' \times Q) \qquad \text{(see 62,63,64,65) (A.27)}$$

---

62 Here we use the property: $Cov(\sum_i^n X_i, \sum_j^m Y_j) = \sum_i^n \sum_j^m [Cov(X_i, Y_j)]$.

63 Here we use the property: $Cov(a \times X, b \times Y) = a \times b \times Cov(X, Y)$.

64 Here we use the property: $\sum_j^n \sum_j^n (X_j) = \sum_j^n (X_j)$.

Therefore, when taking expectation, Eq. (A.25) can be rewritten as follows:

$$E_0[\left.\frac{\partial^2 \ln L}{\partial\lambda\partial\sigma_H^2}\right|_1] = -\ \frac{1}{\sigma_H^2} \times trace\{[(I - \boldsymbol{\lambda} \times W)^{-1}]' \times W'\} \tag{A.28}$$

The second term of Eq. (A.10), $\left.\frac{\partial^2 \ln L}{\partial\lambda\partial\sigma_H^2}\right|_2$, is non-stochastic, hence, we have:

$$E_0[\left.\frac{\partial^2 \ln L}{\partial\lambda\partial\sigma_H^2}\right|_2] = -\ \frac{\sigma_C^2 - \sigma_H^2 \times \frac{H'\times H}{n} \times Tr}{\sigma_C^4} \times \frac{H'\times H}{n} \times \boldsymbol{Tr_\lambda} \tag{A.29}$$

In addition, by Eq. (A.17), the third term of Eq. (A.10), $\left.\frac{\partial^2 \ln L}{\partial\lambda\partial\sigma_H^2}\right|_3$, when taking expectation, can be expressed as follows:

$$E_0[\left.\frac{\partial^2 \ln L}{\partial\lambda\partial\sigma_H^2}\right|_3] = \frac{\sigma_C^2 - 2\times\sigma_H^2 \times \frac{H'\times H}{n} \times Tr}{\sigma_C^4} \times \frac{H'\times H}{n} \times \boldsymbol{Tr_\lambda} \tag{A.30}$$

Adding up Eqs. (A.28) to (A.30), we finally have:

$$E_0[\frac{\partial^2 \ln L}{\partial\lambda\partial\sigma_H^2}] = E_0[\frac{\partial^2 \ln L}{\partial\sigma_H^2\partial\lambda}] = -\frac{1}{\sigma_H^2} \times trace\{[(I - \boldsymbol{\lambda} \times W)^{-1}]' \times W'$$

$$-\frac{\sigma_H^2 \times \frac{H'\times H}{n} \times Tr}{\sigma_C^4} \times \frac{H'\times H}{n} \times \boldsymbol{Tr_\lambda} \tag{A.31}$$

Now we have all the information we need to form the information matrix, which we denote as $I(\theta_0)$. Substituting Eqs. (A.11), (A.20), (A.22) to (A.24) and (A.31) into Eq. (4.53), we have:

---

65 Here we use the property: $trace(A'\times B) = \sum_i^n \sum_j^n (a_{ij} \times b_{ij})$.

$$I(\theta_0) = \begin{bmatrix} -E_0[\frac{\partial^2 \ln L}{\partial \beta^2}] & -E_0[\frac{\partial^2 \ln L}{\partial \beta \partial \lambda}] & -E_0[\frac{\partial^2 \ln L}{\partial \beta \partial \sigma_H^2}] \\ -E_0[\frac{\partial^2 \ln L}{\partial \lambda \partial \beta}] & -E_0[\frac{\partial^2 \ln L}{\partial \lambda^2}] & -E_0[\frac{\partial^2 \ln L}{\partial \lambda \partial \sigma_H^2}] \\ -E_0[\frac{\partial^2 \ln L}{\partial \sigma_H^2 \partial \beta}] & -E_0[\frac{\partial^2 \ln L}{\partial \sigma_H^2 \partial \lambda}] & -E_0[\frac{\partial^2 \ln L}{\partial (\sigma_H^2)^2}] \end{bmatrix} \tag{A.32}$$

Q.E.D.

CURRICULUM VITAE

Lu Liu
Department of Applied Economics
Utah State University

Current Position

Graduate Research Assistant (Jul, 2007- Jul, 2010)

Education

PhD: Utah State University, Department of Applied Economics
Logan, Utah, USA, Aug, 2006 - Aug, 2010, Expected Graduation: Jul, 2010
Major: Environmental and Resource Economics, Cumulative GPA: 3.89

Dissertation Title: Three Essays on Environmental and Spatial Based Valuation of Urban Land and Housing

Major Professor and Committee Members: Dr. Paul Jakus (major professor), Dr. Arthur Caplan, Dr. Reza Oladi, Dr. Frank Caliendo and Dr. Mark Brunson

MA: Southwestern University of Finance and Economics, School of Economics
Chengdu, Sichuan, China PR, Sep, 2003 - Jun,2006, Major: Regional Economics

BA: Southwestern University of Finance and Economics, School of Economics
Chengdu, Sichuan, China PR, Sep, 1999 - Jun, 2003
Major: Economics (National Training Center for Experts in Economics)

Professional Experience

Project Manager, Chengdu SAGA Organization Ltd.., Chengdu, Sichuan, China PR, Feb, 2006 - Jun, 2006

Land and Housing Market Analyst, Chengdu SAGA Organization Ltd.., Chengdu, Sichuan, China PR, Sep, 2004 - Jan, 2006

Research Interests

Environmental and Resource Economics & Regional Economics

I am interested in general environmental and resource issues, especially in the application to land and housing markets, as well as other regional economic issues. My research has used discrete choice modeling, spatial, and time series econometrics. In addition, I am interested in the use of general equilibrium approaches to the study of Environmental and Resource Economics & Regional Economics under both autarky and open economy settings.

Publications and in Review

Jakus, P. M., Keith, J. E., **Liu**, **L.,** Blahna, D., 2010. The welfare effects of restricting off-highway vehicle access to public lands. Agricultural and Resource Economics Review 39 (1) (February), 89–100.

Abstract

Off highway vehicle (OHV) use is a rapidly growing outdoor activity that results in a host of environmental and management problems. Federal agencies have been directed to develop travel management plans to improve recreation experiences, reduce social conflicts and diminish environmental impacts of OHVs. We examine the effect of land access restrictions on the welfare of OHV enthusiasts in Utah using Murdock's (2006) unobserved heterogeneity version of the random utility model. Our models indicate that changing access to public lands from fully "open" to "limited" results in relatively small welfare losses, but that prohibiting access results in much larger welfare losses.

Liu, **L.**, Oladi, R., 2009. Pollution permits, abatement, and international trade. In review.

Abstract

This paper studies the optimal pollution emission and abatement policies comprehensively for a small economy. We present a dynamic general equilibrium model of pollution, emission permit, abatement, and international trade. With free trade the government employs a two-dimensional policy. On the one hand, the government issues pollution permits and therefore controls the emission level. On the other hand, it undertakes a positive level of abatement activities. In contrast, the government is inactive for the latter dimension of the policy under the autarky equilibrium.

Jakus, P. M., Keith, J. E., **Liu**, **L.**, 2008. Economic impacts of land use restrictions on OHV recreation in Utah: a report for the Utah governor's Public Lands Policy Coordination Office. Available at: http://governor.utah.gov/publiclands/PLPCOStudies/OHV_EIC_2008.pdf

Presentations

Liu, **L.**, 2009. A spatial hedonic study for monopoly supplied urban land via English auction: a case study of Chengdu, China. Presentation at Department of Applied Economics Seminar Series, Department of Applied Economics, Utah State University, Logan, UT, USA (November).

Liu, **L.**, Oladi, R., 2009. Pollution permits, abatement, and international trade. Presentation at Department of Applied Economics Seminar Series, Utah State University, Logan, UT, USA (November).

Liu, **L.**, Oladi, R., 2009. Pollution permits, abatement, and international trade. Selected presentation at International Economics and Finance Society China (IEFS China) Inaugural Conference, Beijing, China (May).

Liu, **L.**, 2009. Three models on natural disaster: economic impact, rescue and reconstruction. Selected presentation at 2009 Intermountain Research Symposium, Logan, UT, USA (April).

Liu, **L.**, 2008. The relationship between housing mortgage rate and income tax rate: a general equilibrium with housing and public good. Selected presentation at Hong Kong Economic Association (HKEA) Fifth Biennial Conference, Chengdu, China (December).

Liu, **L.**, 2008. The relationship between housing mortgage rate and income tax rate: a general equilibrium with housing and public good. Selected presentation at 2008 Intermountain Symposium, Logan, UT, USA (April).

Liu, **L.**, 2007. A quantitative method for optimal decision in urban residential development. Selected presentation at 2007 China Economic Annual

Conference (CEAC), Shenzhen, China (December).

Academic Memberships

American Economic Association (Student Member).
Western Agricultural Economics Association (Student Member).
International Economics and Finance Society China (Student Member).
Global Chinese Real Estate Congress (Student Member).
Goldenkey.

Community Service

Department Representative in the Graduate Student Senate, 2008-2009.

Honors and Awards

Third Place in Presentation (Economics and Business Session), Intermountain Graduate Research Symposium, 2010.
Graduate Student Senate (GSS) Enhancement Award, 2009.
Graduate Student Senate (GSS) Conference Travel Awards, 2008.
Graduate Research Assistantship, 2007-2010.
Vice President for Research Fellowship, 2006.