Love Thy Neighbor: Income Distribution and Housing Preferences

Tin Cheuk Leung
Chinese University of Hong Kong

Kwok Ping Tsang
Virginia Tech*

May 26, 2010

Preliminary - please do not cite

Abstract

Do homeowners prefer living in an area with a more equal distribution of income? We answer this question by estimating a nonparametric hedonic pricing model for about 90,000 housing units transacted in Hong Kong between 2005 and 2006. We first identify a hedonic price function by locally regressing the rental price of the housing unit on its intrinsic and neighborhood characteristics, one of which is the Gini coefficient for household income of the constituency area. We then combine the estimates with a log utility function to obtain the heterogeneous preference parameters. Finally, we estimate the joint distribution of the preference parameters and demographics. We find that most homeowners have a strong distaste for inequality in their neighborhood, and the distaste increases with income and goes down with age. Counterfactual experiments show that reallocating Public Rental Housing by half can increase the welfare of homeowners by about HK$8,000 on average per year, which is equivalent to increasing the housing unit by 20 square feet or reducing the age of the unit by 5 years.

*First draft: March 2010. Leung: Department of Economics, Chinese University of Hong Kong, Room 914, Esther Lee Building, Chung Chi Campus, Shatin, Hong Kong. Email: tleung@cuhk.edu.hk. Tsang : Department of Economics, Virginia Tech, Pamplin Hall (0316), Blacksburg, Virginia, 24061. Email: byront@vt.edu.

1 Introduction

"How seldom we weigh our neighbor in the same balance with ourselves"

Of the Imitation of Christ, Thomas à Kempis (1418)

Do homeowners have a preference for living among neighbors with a similar income level? Common sense suggests that homeowners prefer income equality in the neighborhood. Despite the alleged snobbery of the rich towards the poor and the reciprocal jealousy of the poor towards the rich, people with a similar income living in the same area also creates positive externality: more goods and services are available catering for an income group if more people in that income group are living in the same area. For example, a low-income household will have a hard time finding a cheap fast food chain in a high-income area, and it would be difficult for a high-income household to find an expensive restaurant in a poor neighborhood. Recent neural science research has also shown that humans have social preferences to reduce inequality in outcome distributions (see Tricomi, Rangel, Camerer, and O'Doherty (2010)).

In this paper we give a quantitative answer to this question by studying over 90,000 transactions in the Hong Kong housing market in 2005 and 2006. We first present a simple model that takes the preference of a homeowner as given, and looks at the equilibrium in the housing market when the homeowner prefers to live near others with a similar income level. Specifically, the model tells us how public housing affects the equilibrium distribution of the housing units and the welfare change. Since we are interested only in the implications of the distaste towards income inequality, our model is an *ad hoc* one that abstracts from other important aspects of the housing market. We then describe the data, and explain why two unique features of the Hong Kong housing market are important for our purpose. First, Hong Kong is a densely populated area that magnifies the impact of a neighborhood effect. Second, the public housing policy in Hong Kong has created substantial income inequality within small areas. Using a 3-step nonparametric hedonic pricing technique, we obtain the willingness to pay and preference parameters for the characteristics the housing unit and its neighborhood. In particular, we look at the preference for income inequality and see how the preference changes with the demographics. Finally, we conduct a counterfactual experiment by reallocating half of the poorest public housing residents in all constituency areas in Hong Kong, and look at the welfare implications.

To the best of our knowledge, this is the first paper that estimates the preference of residents for income inequality in the neighborhood. There is some empirical research on the neighborhood effect or housing externalities among residents. When there is an urban renewal project for some area, does the land value of the nearby area also increase? Using the American Housing Survey for 1985 and 1989, Ioannides (2002) finds that whether the neighbors (the 10 nearest housing units) of an individual have house maintenance substantially affects the individual's maintenance decision. That is, living in a dilapidated neighborhood discourages an individual to improve her housing unit, and the individual has a higher incentive to renovate when the neighbors' housing units look much better. A recent paper by Rossi-Hansberg, Sarte, and Owens (2010) looks at the concentrated residential urban revitalization programs in Richmond, VA. A few disadvantaged neighborhoods (the impact area) are supported by the federal government to renovate, but with the neighborhood effects the neighborhood of the impact area also benefits from the program. The authors find that there is an increase in the land value of the neighborhood, and the effect decreases with the distance from the impact area. Our paper differs from the housing externalities literature that we identify the *preference* for a neighborhood in a hedonic pricing framework instead of the *behavior* induced by the neighborhood. We quantify the distaste of people for income inequality in the neighborhood, and we measure the welfare gain of reducing income inequality in the residential areas in Hong Kong.

This paper is also related to the literature of income sorting in residential areas, at least since Tiebout (1956). A local government collects taxes and provide public goods, and communities are formed endogenously. The Tiebout model predicts income stratification across communities, or all people in each community have the same marginal benefit from the local public goods. Allowing for heterogenous preferences, Epple and Platt (1998) show that it is possible reduce the amount of stratification, i.e. people with the same income are not necessarily in the same community. Using data from the American Housing Survey, Ioannides (2004) finds that within small neighborhoods (same as Ioannides (2002)), there is substantial income mixing. In our empirical framework each homeowner takes the income distribution in the constituency area as given, and from the data we infer how much the homeowner is willing to pay for less income inequality in the area. Income distribution within constituency areas in Hong Kong are not only determined by the competitive market, but are influenced by public housing policy (see Section 2). Based on the estimated preference, we look at the welfare implications of reducing income inequality in the constituency areas.

2 Why the Hong Kong Housing Market?

Hong Kong is famous for being a densely populated city. According to the World Population Prospects (2008 revision), the estimate of population density in Hong Kong is 6,433 people per squared kilometre for the year 2010, in contrast with 33 people in the United States, 225 people in the United Kingdom and 336 people in Japan. Of course, Hong Kong is less populated than major cities in the US like Manhattan, New York (25,850 people, Census Bureau data). But since Hong Kong is characterized by being mountainous and high-rises, some of the residential areas we study can be highly populated. For example, a medium-quality high-rise of 40 floors usually have more than 16 housing units on each floor, and residents are forced into having frequent interactions with neighbors (in the elevator, or even hearing a conversation from next door). As a result, the existence of a neighborhood effect will be the most convincing for the case of Hong Kong than in the areas studied in the literature.

As our interest is the preference for income inequality in the neighborhood and the potential benefit of removing inequality, we need to study residential areas with significant variations of income distribution. The public housing policy in Hong Kong has contributed to the substantial income inequality in different areas of Hong Kong.

In 1953, a fire in Shek Kip Mei destroyed thousands of shanty homes. Since then, the government of Hong Kong began to construct homes for the poor. A significant portion of people in Hong Kong are inhabiting in public housing. According to 2006 census, 3.4 million people, out of 6.9 million, lived in public housing provided by the Hong Kong government. This is the greatest government intervention in a city renowned for its free-market principle.

There are three main types of public housing in Hong Kong. The first type is Public Rental Housing estates which are the most numerous type of public housing. As of 2006, 2.1 million people lived in Public Rental Housing estates. Applicants' income and total net assets value cannot exceed certain limits, which vary between families, the elderly and individual applicants.

For instance, the monthly income and total net asset limit for a two-person household are HK$11,660 and HK$252,000.

The second type is the Home Ownership Scheme estates. These are subsidized-sale public housing estates for low-income residents. As of 2006, 1.2 million people lived in these estates. The income and asset limits are higher than that of the Public Rental Housing estates. The monthly income and total net asset limit for a two-person household are HK$23,000 and HK$660,000.

The third type is the Sandwich Class Housing Scheme estates. They were built for sale to the "sandwich class", which are the lower-middle and middle-income residents not eligible for other public housing but have difficulties affording private housing. The flats are sold at prices slightly below market value (usually 70%), but with quality comparable to some middle-class private housing. The supply of these estates are limited. Only 48,106 people lived in these Sandwich Class Housing Scheme estates in 2006.

For purpose of elections in the District Council, the government divide Hong Kong into 18 districts. In each districts, there are several sub-districts (constituency areas). In total, there are 380 constituency area.

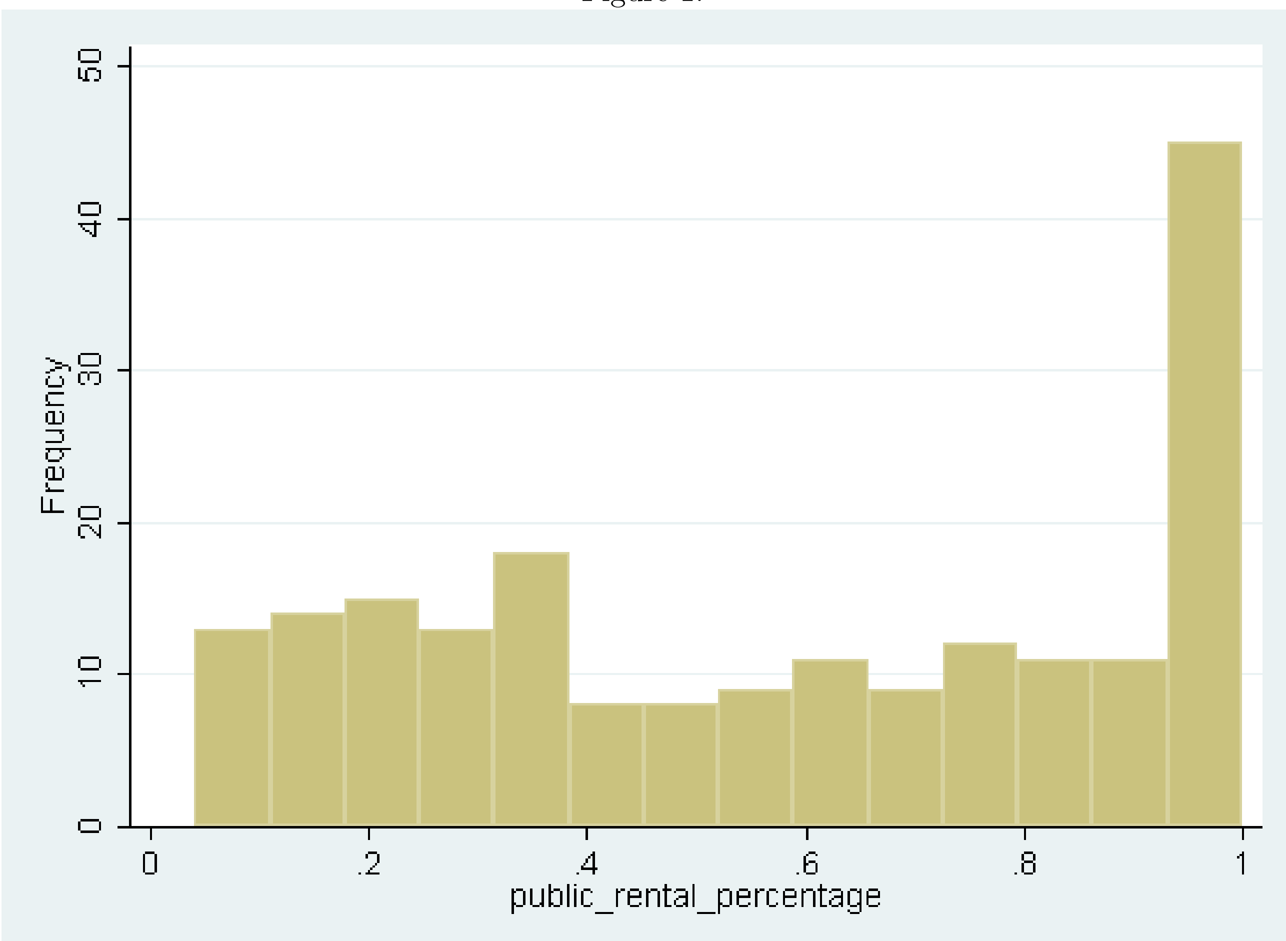
The Census data only allows us to separately identify the demographics of the Public Rental Housing estates and the rest. Thus, in the analysis below, we focus to analyze on how Public Rental Housing estates affect Gini and welfare of homeowners in different constituency area.

One distinct feature in the Hong Kong housing market is that public housing inhabited by lower-income group and private housing inhabited by higher-income group can coexist in the same neighborhood (or the same constituency area). While about half of the constituency areas (186 out of 380) do not have any Public Rental Housing estates, Figure 1 shows that the percentage of Public Rental Housing units in a constituency area varies evenly across the rest of the constituency areas.

3 A Stylized Location Model

Given that people have a distaste for income inequality in the neighborhood, what can we say about the location choice of the people? How does the introduction of public housing affect

Figure 1:



Percentage of Public Rental Housing Across Constituency Areas (excluding constituency areas without any public rental housing) the location choice? Does public housing reduce the welfare of some people? To illustrate the main points of this paper, in this section we present a highly stylized model that takes the preference of residents as given, and abstracts from many other aspects of the housing market.

In the model, there is a unit measure of agents, $i \in [0, 1]$. Each agent is endowed with income $w(i)$, with $w'(i) \geq 0$, so agent 0 is the poorest and agent 1 is the wealthiest. Given $w(i)$ and $w(-i)$, agent i chooses to stay in neighborhood n, where $n = 1, \ldots, N$. Denote the policy function of agent i, which will be discussed further below, as $P(i|-i)$.

In this model, an agent has a distaste to stay in a neighborhood in which neighbors' incomes are very different from that of the agent. The utility of agent i living in neighborhood n is:

$$u(i,n) = -\int_0^1 \frac{P(j|-j)[w(i) - w(j)]^2}{S_n}\, dj$$

where $S_n = \int_0^1 P(j|-j)\, dj$.

Agent i's problem is to choose a neighborhood that maximizes his utility. Thus the policy function solves the fixed point problem below:

$$P(i|-i) = \arg\max_k u(i,k) \tag{1}$$

We have not been able to prove the existence and uniqueness of the fixed point. Thus, no analytical solution can be provided at the moment. Instead, we solve the fixed point numerically and find that the result is fairly robust.

The set up of the numerical exercise is as follows. We have 200 home buyers and 5 locations. Income of each home buyer is drawn from a log-normal distribution. Income is then sorted from low to high, so that $w(1) \leq w(2) \leq \ldots \leq w(200)$. Each home buyer chooses to locate in one of the 5 neighborhoods as modeled above. The solution of the problem is a 200-by-1 vector policy function for the 200 agents. That is, each policy function solves equation (1). To see if the result is robust to different income draws, we repeat the above steps 100 times. The average location and welfare can be interpreted as the expected location and welfare of the home buyer.

We solve two equilibria. First, it is the free market equilibrium in which every agent is free to choose his neighborhood. Second, it is the public housing equilibrium in which the poorest 50 agents are assigned to two neighborhoods.

Figure 2 shows the income distribution in each of the five neighborhoods in the two equilibria. The first column is the free market equilibrium, and the second column is the public housing equilibrium. There are two things to note. First, aside from the richest neighborhood, the incomes in all other neighborhoods are more spread in the public housing equilibrium. This means all but the richest agents would be worse off when there is public housing. Second, the income of agents living in private housing is lower in the two neighborhoods with public housing, which is empirically testable.

4 Data Description

In this paper, we use housing transaction data, provided by Economic Property Research Center, between 2005-06 as our main source of data.[1] We then supplement this data with 2006 Hong Kong by-census data, available on the internet.

4.1 Housing Transaction Data

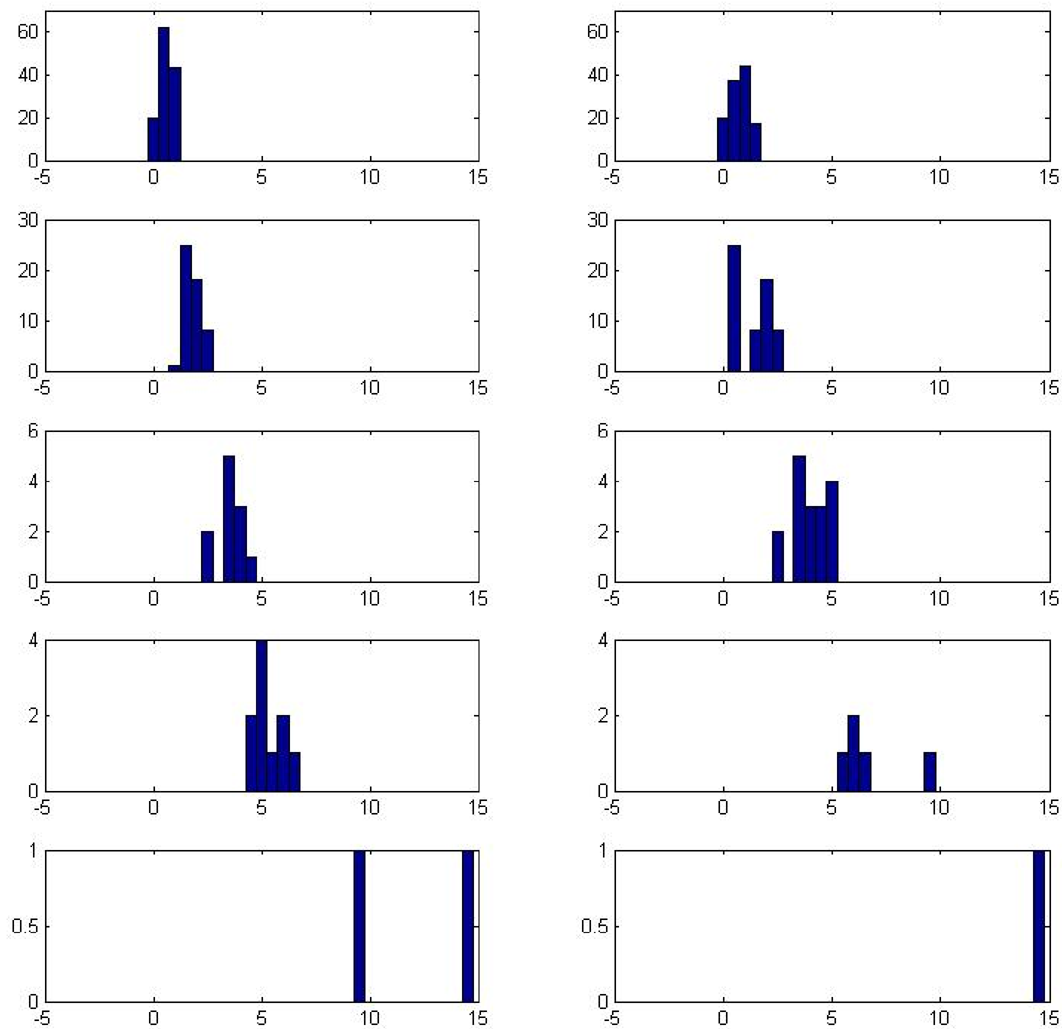
The housing transaction data contains many micro-aspect of each transaction, including prices, gross and net area, address, floor, age, number of bedrooms and living rooms, and so forth.

This data have pros and cons compared with more conventional micro data from US Census of Population and Housing. On the one hand, the census data provide more detailed homeowners' demographic and financial information than our data. We can only use the average demographic information of people living in private housing in various constituency areas to proxy it. But, on the other hand, the home price data from the Census is self reported and is top coded at US$875,000. In addition, home prices are partitioned into only 23 mutually exclusive categories. Also, Census data only provide limited information on the home's characteristics such as the number of rooms and age of the structure.

There are 357,931 observations in the initial data. We drop observations with missing characteristics like prices, floor, area e.t.c. We then select the observations in major estates and building in each constituency area and merge it with census data which has the demographic

[1] For studies on the Hong Kong housing market using the same dataset, see Leung, Lau, and Leong (2002) and Leung, Leong, and Wong (2006).

Figure 2:



Income Distribution in Different Neighborhood (Column 1: free market equilibrium; Column 2: public housing equilibrium)

Table 1: Data Selection in the Homes Transaction Data

Reasons for exclusion	Year 2005 # dropped	Year 2005 # remain	Year 2006 # dropped	Year 2006 # remain
Initial Sample	N.A.	173445	N.A.	184486
Missing Floor	4097	169348	3613	180873
Missing Gross Area	45072	124276	53017	127856
Missing Net Area	14249	110027	15429	112427
Missing Bedroom	26587	83440	27691	84736
Missing Living Room	31	83409	24	84712
Price = 0	1731	81678	2021	82691
Price Outliers	1621	80057	1652	81039
Non-major Estates/Buildings	33790	46267	38216	42823

Table 2: Summary Statistics for Hong Kong Homes Transacted in 2005-06

	Year 2005	Year 2006
Price (\$HK Million)	2.46 (1.81)	2.51 (2.02)
Floor	18.63 (12.77)	18.31 (12.29)
Gross Area (sq ft)	713.61 (248.34)	722.07 (263.45)
Net Area (sq ft)	561.94 (207.10)	572.86 (224.54)
Bedrooms	2.38 (0.55)	2.40 (0.56)
Living Rooms	1.87 (0.33)	1.86 (0.34)
Age of Structure	12.42 (7.85)	14.07 (7.76)
Swimming Pool	0.80 (0.40)	0.78 (0.41)
Club House	0.54 (0.50)	0.52 (0.50)
std. dev. in parenthesis.	N = 46267	N = 42823

statistics for each constituency area. This leaves us with 89090 observations. Table 1 describes data selection process.

Table 2 summarizes the characteristics of the transacted housing units between 2005-2006. The transaction prices and the transacted housing units are reasonably similar across the two years, which can serve as a justification for treating the two years as the same market in the later analysis.

4.2 Census Data

The 2006 by-census data includes various demographic, social, educational, economic and household information of each of the 380 constituency areas. Table 3 summarizes the demographic information of the constituency areas. The average population in a constituency area is 17,148. All of the variables reported in Table 3 have reasonable variation across constituency areas.

Income inequality of a neighborhood is an important attribute for buying a home. Despite

Table 3: Summary Statistics of Constituency Areas

	Mean	Std. Dev.	Min.	Max.
Population	17148.22	4465.66	5158	46447
Median Age	39.72	3.23	26	50
Median Household Income	10769.88	2959.966	7000	25000
Average Household Size	3.00	0.32	2.3	4
% Post Secondary Education	22.25	12.69	5	64.9
Gini	0.45	0.06	0.22	0.60
% Public Rental Housing	0.30	0.37	0	1

N=380

the alleged snobbery of the rich towards the poor and the reciprocal hatred of the poor towards the rich, people with a similar income living in the same area also creates positive externality: more goods and services are available catering for an income group if more people in that income group are living in the area. Also, as Gans (1961) points out, tension may arise from differences in child-rearing norms among different income groups. While "people with higher incomes and more education may feel that they or their children are being harmed by living among less advantaged neighbors. The latter are likely to feel equally negative about the 'airs' being put on by the former.". As shown in Table 3, the gini coefficients across constituency areas vary a lot from 0.22 to 0.6, average at 0.45.

5 Model

In this section, we build a model of housing demand for households in Hong Kong between 2005 and 2006. A home $j = 1, \ldots, J$ is a bundle of three types of characteristics: physical attributes, community attributes and attributes observed by consumer but not by econometricians. Physical attributes include the number of rooms, the age of the unit, gross and net area of the unit, number of bedrooms and living rooms, and dummy variables for swimming pool and club house. Community attributes include average demographics of persons in the constituency area, such as the fraction of college educated households. The unobserved attribute is modeled as a scalar ξ_j.

Prices of houses are determined by the interaction of buyers and sellers in the equilibrium. The price function $\mathbf{p}$ maps housing characteristics $(\mathbf{x}, \xi)$ into their equilibrium prices:

$$p_j = \mathbf{p}(\mathbf{x}_j, \xi_j) \tag{2}$$

Households take prices as given and solve the following static utility maximization problem:

$$u_{ij} = \max_j u_i(\mathbf{x}_j, \xi_j, c) \text{ Subject to: } p_j + c \leq y_i \tag{3}$$

where c is a composite commodity, with a price normalized to \$1 (pre-tax).

Suppose the characteristic k is continuous and that j^* is household i's optimal choice. The first-order condition of equation (3) says that the marginal rate of substitution between product characteristics k and the composite commodity must equal to the implicit price:

$$\frac{\frac{\partial u_i(\mathbf{x}_{j^*}, \xi_{j^*}, y_i - p_{j^*})}{\partial x_{j,k}}}{\frac{\partial u_i(\mathbf{x}_{j^*}, \xi_{j^*}, y_i - p_{j^*})}{\partial c}} = \frac{\partial \mathbf{p}(\mathbf{x}_{j^*}, \xi_{j^*})}{\partial x_{j,k}} \tag{4}$$

As noted by Bajari and Benkard (2005b) and Bajari and Kahn (2008), a single cross section observed in this data is not enough to recover a household's utility function globally. We follow the literature on random coefficient discrete choice models to specify household's utility to be:

$$u_{ij} = \beta_i' [\log(\mathbf{x}_j); \log(\xi_j)] \tag{5}$$

We allow for a rich specification of heterogeneity in tastes as we allow the marginal valuation of the characteristics to be household specific, since β_i are household specific. Also, utility in equation (5) is a log-linear function of the product characteristics. The log specification allows product characteristics to have diminishing marginal utility.

Most of the previous studies on differentiated product assume β_i to have a parametric distribution. In particular, they are independently and normally distributed.[2] We do not impose any parametric distribution on β_i and will estimate the distribution of new homeowners' tastes nonparametrically.

In this paper, we are interested to see how distaste against income inequalities of households with different demographic characteristics differ. We thus model the joint distribution of the

[2] See Berry, Levinsohn, and Pakes (1995), Petrin (2002), Nevo (2000), and Rossi, Allenby, and McCulloch (2005).

random utility coefficients, β_i, and demographics.[3] As discussed in Bajari and Kahn (2008), the lack of micro level data on household level characteristics requires an assumption of linearity between tastes and demographics.

6 Estimation

Our estimation approach involves three steps. The first two steps are similar to those used in Bajari and Benkard (2005a); the last step is similar to Bajari and Kahn (2008). In the first step, we estimate the hedonic price function $\mathbf{p}$ using a flexible local linear regression method described in Fan and Gijbels (1996) and applied in Bajari and Kahn (2005) and Bajari and Kahn (2008). Second, we "back out" the random utility coefficients for each household by applying first order conditions for optimality. Finally, we recover the joint distribution of random utility coefficients and household demographics. Since we only have access to demographics aggregated at the level of District Councils constituency areas, we follow Bajari and Kahn (2008) to estimate household level preferences with this aggregated data.

6.1 First Step: Estimating the Hedonic Price Function

We follow Fan and Gijbels (1996) to use local linear methods to estimate the hedonic flexibly. For a particular home j^*, we assume the hedonic price function $\mathbf{p}$ is locally linear and satisfies:

$$p_j = \alpha_{0,j^*} + \sum_k \alpha_{k,j^*}(x_{j,k} - x_{j^*,k}) + \xi_j \tag{6}$$

We only assume the hedonic in equation (6) is locally linear, not globally linear as in a linear regression model. The coefficients $\alpha_{\cdot,j^*}$ have a subscript j^* to emphasize that they are specific to (x_{j^*}, ξ_{j^*}).

For any $j^*, 1 \leq j^* \leq J$, we follow Fan and Gijbels (1996) to use weighted least squares to estimate α_{j^*}

[3] Section 7 provides more discussion on the set up.

$$\alpha_{j^*} = \arg\min_{\alpha}(\overrightarrow{\mathbf{p}} - \mathbf{X}\alpha)'\mathbf{W}(\overrightarrow{\mathbf{p}} - \mathbf{X}\alpha) \tag{7}$$

$$\mathbf{p} = [\text{RPRICE}_j], \mathbf{X} = [\mathbf{x}_j], \mathbf{W} = diag\{K_h(\mathbf{x}_j - \mathbf{x}_{j^*})\} \tag{8}$$

In equations (7) and (8), $\overrightarrow{\mathbf{p}}$ is the vector of the owner's equivalent rent for all homes $j = 1, \ldots, J$, $\mathbf{X}$ is a vector of regressors which correspond to the observed product characteristics and $\mathbf{W}$ is a matrix of kernel weights.

The kernel weights in $\mathbf{W}$ are a function of the distance between home j^* and j. The local linear regression assigns more weights to observation near j^*. As discussed in Fan and Gijbels (1996), local linear methods have the same asymptotic variance and a lower asymptotic bias than the Nadaraya-Watson estimator, whereas the Gasser-Mueller estimator has the same asymptotic bias and a higher asymptotic variance than local linear methods. We chose the following normal kernel function with a bandwidth of 3:

$$K(z) = \prod_k N(z_k/\hat{\sigma}^2) \tag{9}$$

$$K_h(z) = = K(z/h)/h \tag{10}$$

In equation (9), K is a product of standard normal density and $\hat{\sigma}^2$ is the standard sample deviation of characteristic k.

We interpret the residual the hedonic regression from equations (7) and (8) as the unobserved home characteristic.

$$\xi_{j^*} = p_{j^*} - \mathbf{x}_{j^*}\alpha_{j^*} \tag{11}$$

In the first step, we run a linear regression of owner equivalent rent on $\mathbf{x}_j$ and constituency area fixed effect. The constituency area fixed effects absorb important attributes such as distance from work, air quality, crime rate and local school quality. We then subtract the constituency area fixed effects from the owner's equivalent rent and estimate the local linear regressions described above.

The treatment for binary variables (e.g. the presence of a swimming pool) is different. Suppose the household i chooses a house j^*. Define $\hat{\mathbf{x}}_j$ as the observed characteristics of house j^* except one of the binary variable x is set to 1, and $\bar{\mathbf{x}}_j$ as the same characteristics with the binary variable set to 0. The implicit price for the binary characteristic x is then $\mathbf{p}(\hat{\mathbf{x}}_j, \xi_j) - \mathbf{p}(\bar{\mathbf{x}}_j, \xi_j)$, and if household i chooses $x = 1$ then $\beta_{i,x} > \mathbf{p}(\hat{\mathbf{x}}_j, \xi_j) - \mathbf{p}(\bar{\mathbf{x}}_j, \xi_j)$ and $\beta_{i,x} < \mathbf{p}(\hat{\mathbf{x}}_j, \xi_j) - \mathbf{p}(\bar{\mathbf{x}}_j, \xi_j)$ otherwise. That is, $\beta_{i,x}$ is not identified.

6.2 Second Step: "Backing Out" the Random Utility Coefficients

Due to the log utility function (5), we can calculate the random utility coefficients easily. Let $\hat{\alpha}_{j^*,k}$ be the estimated coefficients from the local linear regression for variable x_{j^*}. The coefficients are the implicit prices faced by household i, who chooses x_{j^*}, in the market, and hence $\hat{\alpha}_{j^*,k}$ is the estimated implicit price $\frac{\partial \mathbf{p}(\mathbf{x}_{j^*}, \xi_{j^*})}{\partial x_{j,k}}$. The random coefficients for this household i is calculated as:

$$\hat{\beta}_{i,k} = \hat{\alpha}_{j^*,k} x_{j^*,k} \tag{12}$$

That is, we obtain a random coefficient for every characteristic k and for every household i.

6.3 Third Step: Finding the Joint Distribution of Preferences and Demographics

We model the relationship between preferences and demographics using a linear model. Denoting $d_{i,s}$ as the demographic characteristic $s = 1, ..., S$ of household i, we can estimate:

$$\hat{\beta}_{i,k} = \theta_{0,k} + \theta_{k,1} d_{i,1} + \cdots + \theta_{k,S} d_{i,S} + \eta_{i,k} \tag{13}$$

Unfortunately, we do not have observations on the household's characteristics $d_{i,s}$. Instead, we observe the average characteristics of households in each census area $d_{t,s}$ for $t = 1, ..., T$. We follow Bajari and Kahn (2008) and estimate (13) with the group-mean method. We divide the $i = 1, ..., I$ households into G groups each of size $n = I/G$, and write (13) as:

$$\bar{\beta}_{g,k} = \theta_{0,k} + \theta_{k,1} \bar{d}_{g,1} + \cdots + \theta_{k,S} \bar{d}_{g,S} + \bar{\eta}_{g,k} \tag{14}$$

Table 4: Summary of Implicit Hedonic Prices

Variable	Mean	Std. Dev.	25%	50%	75%
Constant	1385.5	95885.4	-59069.1	-14094.1	37535.9
Floor	675.1	62.6	642.2	661.7	691.4
Net Gross Ratio (%)	1274.3	219.4	1167.4	1253.6	1365.7
Gross Area (sq. ft.)	382.1	12.0	375.3	380.5	386.2
Bay Window (sq. ft.)	-290.2	95.7	-317.8	-294.9	-265.6
Age of Structure	-1603.6	146.3	-1682.7	-1639.2	-1570.6
Swimming Pool	36727.7	4095.9	34880.2	35960.6	37460.1
Const. Area Gini	-73166.5	32404.9	-84151.0	-69586.5	-57862.5
Const. Area % Public Housing	-45505.7	6781.4	-48533.5	-46541.0	-44296.8

That is, we regress the mean preference parameter in each group on the mean demographic characteristics of each group. We do not observe these group means either, but we can approximate it by:

$$\bar{d}_{g,s} \simeq \frac{1}{n} \sum_{i \in g} \sum_{t} d_{t,s} \times 1\left\{t(i) = t\right\} \tag{15}$$

The approximation would be close if T and n are large. First, we draw without replacement and group the households into G groups each with n members. Next, we calculate the group average preference $\bar{\beta}_{g,k}$ and average demographics $\bar{d}_{g,s}$ by (15). Third, with the G observations on $\bar{\beta}_{g,k}$ and the G observations on each $\bar{d}_{g,s}$, we can can estimate $\theta_{0,k}, ..., \theta_{k,S}$ for each preference parameter k by OLS.

In Bajari and Kahn (2008) only one draw is made and the OLS standard errors are used for inference. To account for the uncertainty in using group means instead of household-level demographic characteristics, we draw with replacement and estimate (14) for 1000 times. Instead of using the OLS standard errors, we take the standard deviation of the 1000 sets of $\theta_{0,k}, ..., \theta_{k,S}$ estimates to build our confidence intervals.

7 Results and Discussion

7.1 Hedonic Pricing Estimates

In Table 4, we show the results of the first step estimation, hedonic prices for various housing attributes. Since we use a nonparametric regression technique, we display the distribution of the hedonic prices. Most of the average hedonic prices have signs and magnitudes consistent with economic intuition. One floor higher is priced at HK\$675.1 per year. Homeowners would pay, on average, HK\$382.1 per year for each extra square foot in gross area and HK\$1274.3 per year for each percent increase in the net gross ratio. That is, holding the gross area constant, if we increase the net area (for which the home buyer cares more than, say, a bigger swimming pool). Of the community characteristics, home buyers prefer a homogenous neighborhood. Home price drops, on average, by HK\$7,317 when Gini increases by 0.1. In other words, homeowners, on average, are willing to exchange 19 square feet of gross area (about the size of a usual bathroom in Hong Kong) for a decrease in Gini coefficient in the neighborhood by 0.1. Local income equality is a statistically and economically important factor to an average home buyer. In addition, the price of each 1% decrease in the people living in public housing is HK\$455. Home price can drop with more public housing in the constituency area for many reasons: more crime, more traffic or higher population density. Whatever the reason, the presence of this characteristic in the hedonic price function makes sure that the Gini coefficient variable is not measuring any unpleasant effects of public housing, but purely reflecting income distribution.

7.2 Preferences Estimates

In the second step of our third-stage estimation, we use the hedonic price estimates, and his optimal consumption of various housing attributes, to recover his marginal valuation for various housing attributes. In Table 5, we present the distribution of estimates of willingness to pay for a 10% increase in consumption of various attributes. In particular, suppose household i's current consumption of attribute k is x_k, the willingness to pay for an extra 10% for attribute k is:

$$WTP_{i,k} = \beta_{i,k}(\log(1.1x_k) - \log(x_k)) = \beta_{i,k}\log(1.1) \tag{16}$$

Again, most of the estimates have signs and magnitudes consistent with economic intuition. The average homeowner is willing to pay HK\$1,197 per year for home that is 10% higher in floor, HK\$26,348 per year for for a 10% increase in gross area and HK\$9,524 per year for a 10% increase in the net gross ratio. Homeowners are very sensitive to the age of homes. They are, on average, willing to pay almost HK\$2,075 less per year for their homes if they are 10% older.

Table 5: Consumer Willingness to Pay for Housing Attributes

Variable	Mean	25%	50%	75%
Floor	1197.1	542.5	1049.2	1670.6
Net Gross Ratio (%)	9524.0	8644.6	9337.4	10302.2
Gross Area (sq. ft.)	26347.8	19685.8	24120.0	29959.6
Bay Window (sq. ft.)	-599.3	-903.3	-658.4	-225.2
Age of Structure	-2075.4	-2994.7	-2049.2	-1023.0
Swimming Pool	2784.6	3214.6.0	3387.9	3542.2
Const. Area Gini	-3203.6	-3704.7	-3015.0	-2500.5
Const. Area % Public Housing	-205.1	0.0	0.0	0.0

Units are in HK dollars per year.

A more homogenous neighborhood is more preferred. The average homeowner is willing to pay HK$3,204 to avoid the Gini coefficient in the constituency area to increase by 10%. Again, the Gini coefficient is an important consideration for a home buyer: each 1% increase in the Gini coefficient is equivalent to a 1.25% decrease in the size of the housing unit. For example, the home buyer is indifferent between the Gini coefficient going down from 0.50 to 0.45 and the size going up from 1000 square feet to 1012.5 square feet.

Figure 3 plots the first stage coefficients of Gini on the rental price. If we take rental price of the housing unit as a proxy of the buyer's wealth, Figure 3 shows that richer household dislikes income inequality more than less rich household. The third stage estimation described above can enable us to quantify this.

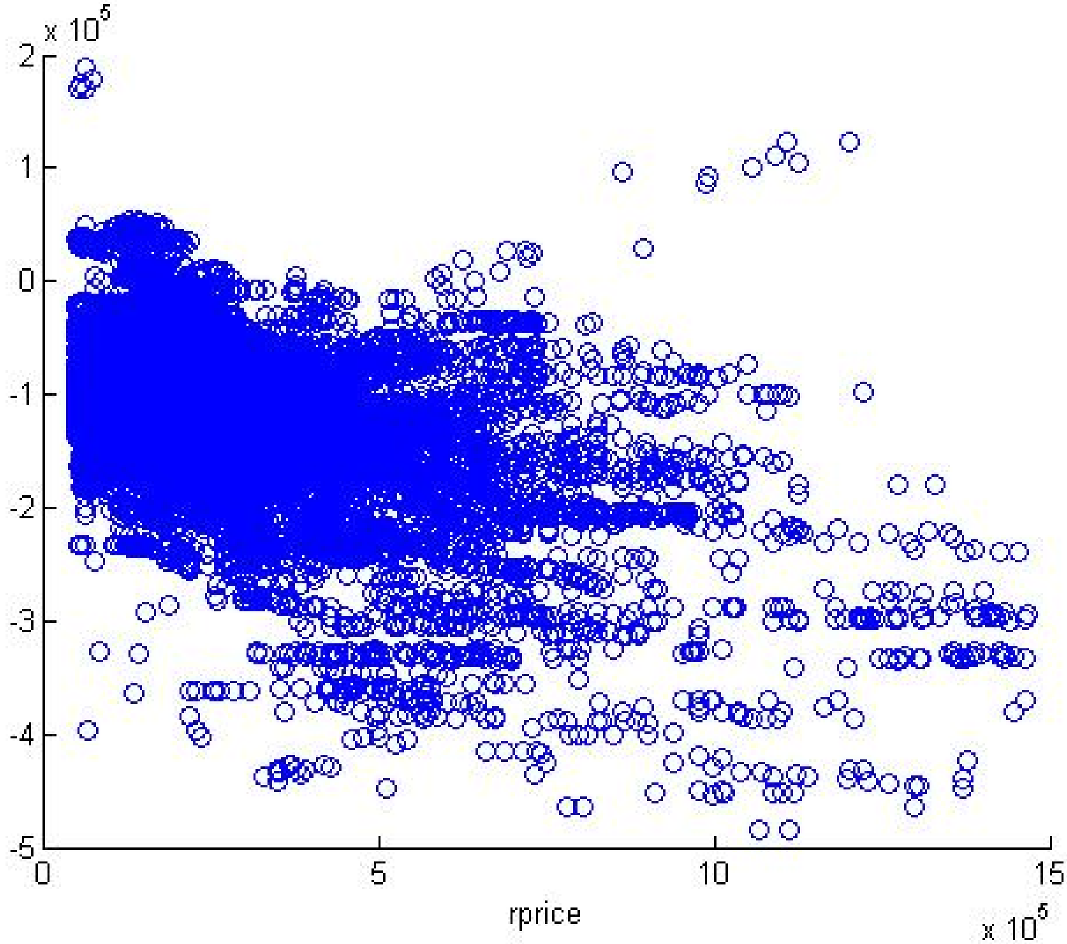
In the third stage estimation, we include four demographic variables in the regression (13):

- age;
- household income ('000);
- marital status (dummies for married, widowed, and separated, and single is the omitted group); and
- education (dummies for less than high secondary, more than high secondary but less than college, and college or above, and high secondary is the omitted group).[4]

For these variables, we exclude the data of people living in public rental housing, who are not homeowners. We then calculate the mean of these variables to be the control variables. In

[4] High secondary means completing Form 5, the level at which a student is about 17 years old. This is roughly equivalent to finishing high school in the US.

Figure 3:



Scatter Plot of Coefficients of Gini and RPrice

Table 6: Willingness to Pay for Income Inequality as a Function of Household Demographics

Age	62.8 (38.4)
Household Income (in $1,000s)	-39.2 (7.0)
% Less than High Secondary Education	-90.6 (23.9)
% Above High Secondary Education	-54.6 (35.2)
% College or Above	-42.2 (19.8)
% Married	35.5 (23.8)
% Widowed	20.2 (67.9)
% Divorced	23.5 (73.3)
% Separated	-192.1 (185.1)
Constant	-1039.5 (2500.2)
Mean R^2– 0.1268	

the calculation of mean age for each constituency area, we exclude certain age groups which are not likely to purchase an apartment. In particular, we exclude people under age of 25. Results are in Table 6. First, Homeowners with higher income dislikes income inequality more. For each HK$1,000 increase of income the the willingness to pay for a 10% increases in Gini goes down by HK$40 per year. Second, older homeowners have a higher tolerance for income inequality. One year increase in age increases the willingness to pay for a 10% higher Gini by HK$62.8 per year. While the distribution of the willingness to pay is weakly related to marital status, it is strongly related to education level. Comparing to the omitted high secondary education group, the lowest education group is much more willing to pay for reducing inequality. For each 1% increase in the % with less than high secondary education, the willingness to pay for a 10% higher Gini increases by HK$90.6 per year. The same holds for the two higher education groups, but by a much smaller amount.

8 Counterfactuals

In previous sections, we show that homeowners have a strong and statistically significant distaste for income inequality in their neighborhood. At the same time, income inequality in some neighborhood is made higher due to the existence of public rental housing. In our data, out of the 89,090 homes transacted between 2005-06, 27,738 of them are located in constituency areas in which there are public rental housing. One natural question to ask is thus: If the Hong Kong government separate private and public housing completely, so that private homeowners do not have public rental housing in their neighborhood, what would be the welfare gain for homeowners?

To answer this question, we do the following counterfactual experiment. We propose the Hong Kong government to reallocate the poorest 50% public rental housing units to constituency areas exclusive to public rental housing. At the same time, we leave the location of homeowners unchanged. This can improve welfare of homeowners through two channels. First, Gini coefficients in some constituency areas decrease. In particular, out of the 199 constituency areas in which we have property transaction data, income inequality in 79 constituency areas changes under this policy. Second, the percentage of public housing units in those 79 constituency area would drop by 50%.

Since most transactions (61,456) took place in constituency areas in which this policy has no effect, the welfare of these homeowners are not affected by this policy. For the rest of the homeowners (27,738), the average welfare gain improves HK$8,126 per year, in which HK$2,150 is due to lower Gini in those constituency areas and HK$5,976 is due to 50% of public housing units in those constituency areas.[5]

Is HK$8,126 per year a large amount? We can get an idea of the magnitude by using our results in Table 5. The amount of welfare gain is roughly equivalent to increasing the housing unit by 20 square feet or having a housing unit 5 years newer. The effect is quantitatively important.

9 Conclusion

People dislike living near others that have a lower or higher income level, and the dislike is substantial: on average, a homeowner is willing to pay about HK$3,200 for a 10% drop of the local Gini coefficient, which is the amount the home buyer is willing to pay for a 1.25% increase in the size of the housing unit. We also find that the dislike of income inequality varies with demographics. It goes up with income and is higher with the low education group. A counterfactual experiment show that relocating part of the Public Rental Housing improves home buyers' welfare. Of course, the experiment ignores the potential problems of grouping all low-income individuals in one area.

[5] We have done the experiment for reducing public housing by 75%, and the welfare change is about HK$12000.

Our results are local, and we ignore how income distribution is endogenously formed in each constituency area. But this paper has identified a large dislike of local income inequality among homeowners which deserves further research. Why do home buyers dislike income inequality, even after (through the fixed effects) controlling for crime, school quality, and the presence of public housing? How should public housing policy be made by taking into account such a preference? How does this preference affect local income distribution over time?

References

Bajari, P., and C. L. Benkard (2005a): "Demand Estimation with Heterogenous Consumers and Unobserved Product Characteristics: A Hedonic Approach," *Journal of Political Economy*, 113(6), 1239–1276.

——— (2005b): "Hedonic Price Indexes with Unobserved Product Characteristics," *Journal of Business and Economics Statistics*, 23(1), 61–75.

Bajari, P., and M. E. Kahn (2005): "Estimating Housing Demand with an Application to Explaining Racial Segregation in Cities," *Journal of Business and Economics Statistics*, 23(1), 20–33.

——— (2008): "Estimating Hedonic Models of Consumer Demand with an Application to Urban Sprawl," *Hedonic Methods in Housing Market*, Springer.

Berry, S., J. Levinsohn, and A. Pakes (1995): "Automobile Prices in Market Equilibrium," *Econometrica*, 63(4), 841–890.

Epple, D., and G. J. Platt (1998): "Equilibrium and Local Redistribution in an Urban Economy when Households Differ in both Preferences and Income," *Journal of Urban Economics*, 43, 23–51.

Fan, J., and I. Gijbels (1996): *Local Polynomial Modelling and Its Applications: Monographs on Statistics and Applied Probability 66.* CRC, London, UK.

Gans, H. J. (1961): "The Balanced Community: Homogeneity or Heterogeneity in Residential Areas?," *Journal of the American Planning Association*, 27(3), 176–184.

Ioannides, Y. M. (2002): "Residential Neighborhood Effects," *Regional Science and Urban Economics*, 32, 145–165.

——— (2004): "Neighborhood Income Distributions," *Journal of Urban Economics*, 56, 435–457.

Leung, C. K. Y., C. K. G. Lau, and Y. C. F. Leong (2002): "Testing Alternative Theories of the Property Price-Trading Volume Correlation," *Journal of Real Estate Research*, 23(3), 253–264.

Leung, C. K. Y., Y. C. F. Leong, and S. K. Wong (2006): "Housing Price Dispersion: An Empirical Investigation," *Journal of Real Estate Finance and Economics*, 32(3), 357–385.

Nevo, A. (2000): "A Practitioner's Guide to Estimation of Random Coefficients Logit Models of Demand," *Journal of Economics and Management Strategy*, 9(4), 513–548.

Petrin, A. (2002): "Quantifying the Benefits of New Products: The Case of the Minivan," *Journal of Political Economy*, 110(4), 705–729.

Rossi, P., G. Allenby, and R. McCulloch (2005): *Bayesian Statistics and Marketing.* John Wiley and Sons, Hoboken, NJ.

Rossi-Hansberg, E., P.-D. Sarte, and R. Owens (2010): "Housing Externalities," *working paper.*

Tiebout, C. M. (1956): "A Pure Theory of Local Expenditures," *Journal of Political Economy*, 64(5), 416–424.

Tricomi, E., A. Rangel, C. F. Camerer, and J. O'Doherty (2010): "Neural Evidence for Inequality-Averse Social Preferences," *Nature*, 463(25), 1089–1092.