Essays on Residential Desegregation

by

Maisy Wong

B.A. Economics
University of California, Berkeley, 2003

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
OCT 1 0 2008
LIBRARIES

Submitted to the Department of Economics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2008

©2008 Maisy Wong. All rights reserved.

The author hereby grants to MIT permission to reproduce and to distribute publicly paper and electronic copies of this thesis document in whole or in part in any medium now known or hereafter created.

Signature of Author..
Department of Economics
August 1, 2008

Certified by..
Esther Duflo
Abdul Latif Jameel Professor of Poverty Alleviation and Development Economics
Thesis Supervisor

Certified by..
Amy Finkelstein
Associate Professor of Economics
Thesis Supervisor

Certified by...... ....
Panle Jia
Assistant Professor of Economics
Thesis Supervisor

Accepted by........... ....
Peter Temin
Elisha Gray II Professor of Economics
Chairman, Departmental Committee on Graduate Studies

ARCHIVES

Essays on Residential Desegregation

by

Maisy Wong

Submitted to the Department of Economics
on August 1, 2008, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Economics

Abstract

Many ethnically diverse countries have policies that encourage integration across ethnic groups. This dissertation investigates the impact and welfare implications of a residential desegregation policy in Singapore, the ethnic housing quotas. I employ both reduced form and structural form methods. In Chapter 1, I use regression discontinuity analysis to estimate the impact of the quota on prices, quantity and quality of units sold. Because an individual's decision on where to locate affects ethnic distributions in aggregate, these externalities suggest that any decentralized equilibrium may not be optimal. To find the first best, just using housing prices is not sufficient because prices do not internalize the externalities. We need to know the shapes of household's preferences. In chapter 2, I use a structural demand estimation framework to estimate taste for living with own ethnic group neighbors. Finally, using these preference estimates, I simulate the first best equilibrium and compare it to the existing equilibrium with quotas.

I find that all quotas have significant negative impact on the proportion of units sold at the quota cutoffs. Malay-constrained units are 5% cheaper perhaps because the units sold are also of lower quality. The impact on the price and quality of Chinese- and Indian-constrained units are opposite. Chinese-constrained units are 7% more expensive even though the units sold are of significantly worse quality. Indian-constrained units are 2% cheaper even though the units sold are of a higher quality.

Using the structural estimation framework in Chapter 2, I find that all groups have strong preferences to live with at least some other members of their ethnic group. However, the shapes of preferences differ significantly across groups. The majority (the Chinese) exhibit preferences that are inverted U-shaped so that after a neighborhood reaches 43% Chinese, they would rather add a new neighbor from the other group. I find similar evidence for the Indians but not for the Malays. My simulations show that the first best has fewer Chinese- and Indian-segregated neighborhoods but more Malay-segregated neighborhoods compared to the existing equilibrium with the quotas. Comparing data from 3 segregated towns before the

quota, I find that after 10 years since the introduction of the quota, the decentralized equilibrium had moved the Malay and Indian proportions significantly closer to first best.

Thesis Supervisor: Esther Duflo
Title: Abdul Latif Jameel Professor of Poverty Alleviation and Development Economics

Thesis Supervisor: Amy Finkelstein
Title: Professor of Economics

Thesis Supervisor: Panle Jia
Title: Assistant Professor of Economics

Acknowledgments

I am deeply grateful to my advisers, Esther Duflo, Amy Finkelstein, Panle Jia and Bill Wheaton for their encouragement, generosity and patience.

Many people have generously offered their time, especially Peter Diamond, Michael Greenstone, Whitney Newey, Patrick Bayer and Stephen L. Ross. I also benefited from conversations with David Autor, Abhijit Banerjee, Kenneth Chay, Glenn Ellison, Sara Ellison, Ben Olken and Robert Townsend. I am grateful to participants of the MIT research lunches in Public Finance, Development, Econometrics, Industrial Organization, Labor and Summer Applied Micro Lunch as well as the MIT Public Finance and Labor seminar and the seminar at the MIT Center for Real Estate.

Graduate school was more manageable because of the support and guidance from Petra Moser, Tavneet Suri and Peter Diamond. I thank Nancy Rose for her support during my job market. I also thank Peet Hoagland, Gary King, Dr. Barbara O'Pray and Katherine Swan for helping me make the transition to MIT.

I acknowledge financial support from the George and O'Bie Shultz Fund and the MIT Graduate Student Fellowship. I thank the Singapore Housing Development Board (HDB) for granting permission to use their data. Much of my work was conducted in the MIT Jameel Poverty Action Lab, with support from Elaine Fulton, Rachel Glennerster, Tricia Gonwa and Courtney Umberger.

My fellow graduate students, Hui Shan, Jessica Cohen, Cynthia Kinnan, Raymond Guiteras, Jim Berry, Greg Fischer and Jeremy Shapiro, greatly enriched my experience at MIT. I would like to especially thank Trang Nguyen for spending hours to talk through my research with me even at the early stages when my ideas were not concrete yet.

I am eternally grateful for the tireless support from my mum, Kam Lee Ching and my brother, Wong Liang Kit, as well as the patience and understanding of my father, Wong Toon Say, his wife, Julie Tan and my sister, Wong Mei Jee. Finally, I wish to thank Najib Wong for his encouragement, his patience and for never giving up on me.

This dissertation is dedicated to the memory of my grandmother, Thong Yoke Pheng.

Contents

Introduction

Many ethnically diverse countries have policies that encourage integration across ethnic groups. In Chicago, the Gautreaux program (the predecessor of the Moving Towards Opportunity program) encouraged residential desegregation using housing vouchers (Jennan, 2000). In the Netherlands, amounts explicitly earmarked for "integration" in the governments' budgets increased from 9 million euros in 1970 to 1.1 billion euros in 2003 (Commissie Blok, 2004). This dissertation investigates the impact and welfare implications of a residential desegregation policy in Singapore, the ethnic housing quotas.

Since Schelling's pioneering work on the theory of segregation, (see Schelling 1969, 1971), the empirical side has received little attention until recently (see Card, Mas and Rothstein, 2008, Bayer, Ferreira and McMillan, 2008). Theoretical models, such as Schelling's, demonstrate the market failure due to externalities. An individual's decision of where to locate affects ethnic distributions, in aggregate. Assuming households care about ethnicities of their neighbors, externalities provide an economic rationale for public policies such as housing quotas. In fact, Schelling (1971) conjectured that ethnic housing quotas could be a useful policy tool to encourage residential desegregation. My dissertation is an empirical investigation of this conjecture.

I employ both reduced form and structural form methods to measure the welfare consequences of the ethnic housing quotas in Singapore. In Chapter 1, I use regression discontinuity analysis to estimate the impact of the quota on prices, quantity and quality of units sold. Externalities suggest that any decentralized equilibrium may not be optimal. To find the first best, we need to know the shapes of household's preferences. In chapter 2, I use a structural demand estimation framework to estimate taste for living with own ethnic group neighbors. Finally, using the preference estimates from chapter 2, I simulate the first best equilibrium and compare it to the existing equilibrium with quotas.

The ethnic housing quotas in Singapore were introduced in 1989 to encourage residential integration amongst the three major ethnic groups in Singapore – Chinese (77%), Malays (14%) and Indians (8%) (Singapore Census, 2000). The policy is a set of upper bounds on block level and neighborhood level ethnic proportions. In practice, the Housing Development Board (HDB) did not want to evict owners in units that were in violation of the quotas. To this day, there exist units above the upper bound. Any transactions that forced these blocks and neighborhoods farther above the upper bound, however, would be barred. For example, when Chinese quotas are binding, non-Chinese sellers cannot sell to Chinese buyers because this transaction increases the Chinese proportion farther above the Chinese quota.

In chapter 1, I use a regression discontinuity framework to estimate the impact of the ethnic quotas on the price, quantity and quality of units sold. I find that all quotas significantly decrease the proportion of units sold at the quota cutoffs. Malay-constrained units are 5% cheaper perhaps because the units sold are also of lower quality. The impact on the price and quality of Chinese- and Indian-constrained units are opposite. Chinese-constrained units are 7% more expensive even though the units sold are of significantly worse quality. Indian-constrained units are 2% cheaper even though the units sold are of a higher quality.

In chapter 2, I estimate the taste for living with own ethnic group neighbors by combining policy induced variation akin to the regression discontinuity identification strategy with a structural demand estimation framework à la Berry, Levinsohn, and Pakes (1995). I find that all groups have strong preferences to live with at least some other members of their ethnic group. However, the shapes of preferences differ significantly across groups. The majority (the Chinese) exhibit preferences that are inverted U-shaped so that after a neighborhood reaches 43% Chinese, they would rather add a new neighbor from the other group. I do not find evidence of such preferences for the Malays.

In chapter 3, I build on previous chapters and answer two questions: What does the first best look like and how close does the quota get us to the first best? I use the preference estimates from the previous chapter to simulate the first best equilibrium. Because of externalities, the decentralized equilibrium is not necessarily the first best. In a competitive equilibrium, the results from chapter 2 suggest that, the average Chinese could have preferences for mixed neighborhoods while the average Malay prefers Malay neighborhoods over mixed neighborhoods. By contrast, the social planner may want to move some Malays into Chinese neighborhoods to have more mixed neighborhoods. While this is not the utility-maximizing choice of individual Malay households, this cost to them has to be balanced against the benefit from providing more diversity for the Chinese. I find that the first best has fewer Chinese- and Indian-segregated neighborhoods but more Malay-segregated neighborhoods. Comparing data from 3 segregated towns before the quota, I find that after 10 years since the introduction of the quota, the decentralized equilibrium had moved the Malay and Indian proportions significantly closer to first best.

Chapter 1

Estimating the Impact of the Ethnic Housing Quotas in Singapore

1.1 Introduction

Many countries have policies that encourage integration across groups of individuals, be it gender groups, ethnic groups or immigrants. In the Netherlands, amounts explicitly earmarked for "integration" in the governments' budgets increased from 9 million euros in 1970 to 1.1 billion euros in 2003 (Commissie Blok, 2004). Many of these desegregation policies take the form of quotas, such as affirmative action quotas for university admissions in Malaysia, hiring quotas in the US police force and quotas on immigrants. This paper studies the impact of a quota policy aimed at residential desegregation in Singapore, the Ethnic Integration Policy.

Since Schelling's pioneering work on the theory of segregation, (see Schelling 1969, 1971), the empirical side has received little attention until recently (see Card, Mas and Rothstein, 2008, Bayer, Ferreira and McMillan, 2008). Theoretical models, such as Schelling's, demonstrate the market failure due to externalities. An individual's decision of where to locate affects ethnic distributions, in aggregate. Assuming households care about ethnicities of their neighbors, externalities provide an economic rationale for public policies such as housing quotas. In fact, Schelling (1971) conjectured that ethnic housing quotas could be a useful policy tool to encourage residential desegregation. This paper is the first step in analyzing this conjecture. I first estimate the impact of the quotas on the housing market in Singapore. In Chapters 2 and 3, I study the welfare implications of the quotas.

The Ethnic Integration Policy was introduced in 1989 to encourage residential integration amongst the three major ethnic groups in Singapore – Chinese (77%), Malays (14%) and Indians (8%) (Singapore Census, 2000). The policy is a set of upper bounds on block level and neighborhood level ethnic proportions. In practice, the Housing Development Board (HDB) did not want to evict owners in flats that were in violation of the quotas. To this day, there exist flats above the upper bound. Any transactions that forced these blocks and neighborhoods farther above the upper bound, however, would be barred. For example, when Chinese quotas are binding, non-Chinese sellers cannot sell to Chinese buyers because this transaction increases the Chinese proportion farther above the Chinese quota. In this way, the quotas work very much like price discrimination. When the Chinese quota binds, the quotas prevent non-Chinese sellers from arbitraging price differences across Chinese and non-Chinese buyers. Thus, prices are allowed to differ across Chinese and non-Chinese buyers when the Chinese quota binds.

What is the impact of the quotas on the housing market? A priori, one may expect no difference in characteristics between constrained and unconstrained units that are very close to the quota. If a unit is $\varepsilon\%$ above the quota (constrained), would buyers be willing to pay a higher price for those units? Why not buy similar units that are $\varepsilon\%$ below the quota and cheaper? Below, I provide a theoretical framework to understand the implications of the quota. Then, I use a regression discontinuity design to test for discontinuities in the price, quantity and quality of units sold, at the quota.

One limitation of studies on desegregation policies is a lack of information on ethnic proportions at a very fine geographic level. To obtain ethnic proportions at the apartment block level, I hand-matched 589,000 names to ethnicities using the Singapore Residential Phonebook. I combined this data with data on monthly transactions downloaded from the HDB website. Finally, I purchased data on the characteristics of apartment blocks from the HDB.

I find that all quotas significantly decrease the proportion of units sold. Malay-constrained units are 5% cheaper perhaps because the units sold are also of lower quality. The impact on the price and quality of Chinese- and Indian-constrained units are opposite. Chinese-constrained units are 7% more expensive even though the units sold are of significantly worse quality. Indian-constrained units are 2% cheaper even though the units sold are of a higher quality.

Below, I discuss the background of the policy (Section 2), provide a theoretical framework to analyze the impact of the quota (Section 3), describe the data (Section 4), discuss estimation (Section 5) and the results (Section 6).

1.2 Background

Singapore is a multi-ethnic country with a population of 4.5 million (Singapore Department of Statistics, 2006). The three major ethnic groups are the Chinese (77%), the Malays (14%) and the Indians (8%). These three groups are very different along many dimensions. The Chinese have the highest median monthly income (S$2335), followed by the Indians (S$2167) and the Malays (S$1790). Although the median Malay household is poorest, the income distribution of the Indians have a longer left tail (more Indians are very poor). Also, the ownership rate in public housing is the lowest amongst the Indians because most of them are renters. The household size is also very different with Chinese families being the smallest, on average (Singapore Census, 2000). Forty-three percent of Malay households have 5 or more family members while only 24% and 26% of Chinese and Indian households have such large families (HDB, 2000)

Public housing is the most popular choice of housing in Singapore with 82% of the resident population living in public housing (HDB, 2006). The flats are built and managed by the Housing Development Board (HDB). There are three ways Singapore residents can live in a HDB flat. They may apply through the primary allocation system for new HDB flats, they may purchase existing HDB flats in the resale market or they may rent. The rental market is negligible: 98% percent of the HDB flats are owner-occupied (HDB, 2006). This paper focuses on the resale market which is the relevant market for the ethnic quotas. Relative to the primary market which is heavily regulated, the resale market functions as an open market.

Public housing was first built in Singapore in 1960 to solve the young nation's housing crisis (HDB website). To cater to the different needs of households, HDB designed and built 8 flat types. Type 1 was a studio, Type 2 meant a 1-bedroom flat, Type 3 was a 2-bedroom flat. Types 4 to 6 all have 3 bedrooms, but the higher types have extra living and/or dining areas. The remainder 2 types are called HUDC and multi-generation units. These tend to be larger units but HDB built very few of them. The most popular flats are type 3 to 6. Apart from the number of rooms, the layout and size in public housing flats are pretty homogenous.

To understand the ethnic quotas, it is important to understand the geography of housing markets in Singapore. The smallest spatial unit is an HDB *flat* A group of HDB flats constitute an HDB *block* A group of HDB blocks make up a *neighborhood.* An HDB block is comparable to a US Census block group, with an average of 70 households. An HDB neighborhood is comparable to a US Census tract, comprising an average of 60 HDB blocks.

The government of Singapore introduced the Ethnic Integration Policy to address the "problem" of the increase in the "concentrations of racial groups" in HDB estates (Parliamentary debate, 1989). The policy was announced in a parliamentary debate on February 16, 1989 and implemented starting March 1, 1989. The Policy is a set of quotas at the block and neighborhood level. Table 1.1 lists the quotas, in comparison to the 2000 national ethnic proportions. Neighborhood quotas are 2% to 8% above national ethnic proportions. Block quotas are 3% above neighborhood quotas, allowing more flexibility at the block level (blocks can be more segregated than neighborhoods). Notice that the Indian neighborhood quotas are only 2% above the national proportion.[1] Consequently, the Indian quotas are more likely to bind. In my sample, 11% of observations have Chinese- and Malay-constrained units but 21% of the observations are Indian-constrained. In practice, the HDB did not want to evict owners in existing flats that were in violation of the quotas. To this day, there exist blocks and neighborhoods above the quota.

The quotas are upper bounds on ethnic proportions to *prevent* HDB communities that are already segregated from becoming more segregated. Once a community hits the upper bound, transactions that make the community more segregated will be blocked. However, other transactions will be allowed. In particular, transactions involving buyers and sellers from the same ethnicity will always be allowed because this does not increase the ethnic proportion. As an example, Table 1.1 shows that the Chinese block quota is set at 87%. Once Chinese make up more than 87% of the block, Chinese buyers can no longer buy from non-Chinese sellers because this increases the proportion of Chinese in that block farther above the quota. Table 1.2 lists the types of transactions barred by each ethnic quota. The important thing to remember is that the ethnicity of the buyer and seller is key. Once a Chinese quota binds, Chinese buyers and non-Chinese sellers are the ones affected. Similarly for Malay and Indian quotas. This restriction prevents arbitrage and thus allows prices

[1] In the parliamentary debates (1989), it was suggested that the quotas were chosen based on the ethnic proportions of buyers in the resale market. Whilst the Indians make up 8% of the population, very few were active in the resale market. Many Indians in public housing were either renters or buyers in the primary market. This is why the percent of Indians allowed at the neighborhood level is only 2% higher than the current national proportion of Indians.

to differ across groups in equilibrium, very much like price discrimination.

1.3 Theoretical Framework

In this section, I adopt the hedonic framework from Rosen (1974) to illustrate how the policy works and how it could impact prices at the quota. I first set up the hedonic framework for the case without quotas, then, I consider what happens with Chinese quotas. The cases for Malay and Indian quotas are similar.

Buyers

Each buyer purchases one unit of a product with n characteristics $(z_1, ..., z_n)$. Write the utility function as $U(x, z_1, ..., z_n)$ assumed strictly concave, where x is all other goods consumed. Set the price of x equal to 1 and measure income, y, in terms of x. The buyer chooses x and $(z_1, ..., z_n)$ to maximize utility subject to a nonlinear budget constraint, $y = x + p(z)$. Defind a bid function, $\theta(z; u, y)$ according to

$$U^i(y - \theta, z_1, ..., z_n) = u \tag{1.1}$$

The bid function, $\theta(.)$ represents the expenditure the buyer is willing to pay for different models of the product (alternative values of z's), given income y, to achieve utility u.

Figure 1.1a illustrates some of these "indifference surfaces" that have been projected onto the $\theta - z_1$ plane cut at the optimal chosen values of the other characteristics, $(z_2^*, ..., z_n^*)$. The curve, $\theta^C(u^*)$, is a Chinese buyer's bid function. It represents the Chinese buyer's willingness to pay for alternative values of z_1, given income y, to achieve utility u^*, holding all other characteristics at the optimum values (For simplicity, the bid function, $\theta^C(z_1, z_2^*, ..., z_n^*; u^*, y)$ is abbreviated as $\theta^C(u^*)$ in the figure). Each buyer has a family of indifference surfaces. For example, $\theta^C(u^*)$ and $\theta^C(u')$ are 2 bid functions for the same Chinese buyer, corresponding to utility u^* and u' respectively, where $u^* > u'$ (a higher bid function corresponds to lower utility because the buyer has to pay more for the same level of z_1). $\theta^{NC}(u^*)$ represents a non-Chinese buyer's bid function. The hedonic price function is the darker line, $p(z_1, z_2^*, ..., z_n^*)$. It represents the minimum price a buyer must pay in the market. The consumer picks the highest utility subject to the non-linear budget constraint, $y = x + p(z)$. One dimension of consumer equilibrium is illustrated in Figure 1.1a, where $\theta^C(u^*)$ is tangent to $p(z)$.[2]

[2]The tangency between the bid function, $\theta(.)$, and the hedonic price function, $p(z)$, is some-

Sellers

The case for the seller is analogous. A seller wants to sell 1 unit of a product. He chooses the characteristics $(z_1, ..., z_n)$ to maximize profits, $\pi = p(z) - c^i(z, \beta)$, where $c(.)$ is a strictly convex cost function with the cost-shift parameter, β representing variables in the cost minimization problem, such as factor prices and production function parameters. Define an offer funcion, $\phi(z_1, ..., z_n; \pi, \beta)$, to be the price a seller is willing to accept for various models of the product, given cost parameter β, to achieve profit π.

Figure 1.1b illustrates a few of these "isoprofit surfaces" projected onto the $\phi - z_1$ plane. $\phi^{NC}(\pi^*)$ and $\phi^{NC}(\pi')$ are two members of a family of these surfaces for a non-Chinese seller where $\pi^* > \pi'$ because a lower offer function means the seller accepts a lower price for the same level of z_1. $\phi^C(\pi^*)$ is an offer function for a Chinese seller. The darker line, $p(z)$, is the maximum price obtainable for each model (each combination of z's) in the market. One producer equilibrium is illustrated in Figure 1.1b where $\phi^{NC}(\pi^*)$ is tangent to $p(z)$ (this is the highest profit attainable given the market prices).

Market equilibrium

Figure 1.1c illustrates a market equilibrium where buyers and sellers are perfectly matched. Each point on the hedonic price function, $p(z)$, is both tangent to the offer function and the bid function. In the case shown, a non-Chinese seller is matched to a Chinese buyer ($\phi^{NC}(\pi^*)$ is tangent to $\theta^C(u^*)$) at z_1^*.

With Chinese Block Quotas

Let z_1 represent the percent of Chinese in a block and let z_1^* be 87% (the Chinese block quota). Figure 1.2 illustrates what happens with Chinese block quotas. Units with z_1 to the left of the vertical line are unconstrained and units with z_1 to the right of the vertical line are constrained. Recall when the Chinese quota binds, non-Chinese sellers cannot sell to Chinese buyers. In this case, the match between the grey offer function and the grey bid function are not allowed ($\phi^{NC}(\pi^*)$ and $\theta^C(u^*)$).

To understand the ethnic housing quotas, the ethnicity of the buyer and the seller are crucial. When the Chinese quota binds, the Chinese buyer and the non-Chinese seller are both affected. Figure 1.2a shows what happens to a Chinese buyer who wants to buy a Chinese-constrained unit. Figure 1.2b shows what happens to a non-Chinese seller who owns a Chinese-constrained unit. Figure 1.2c combines both pictures.

---

what analogous to the tangency between the indifference curve and the budget constraint in the homogenous goods case. Except, here, the budget constraint is non-linear.

Consider a Chinese buyer (θ^C) who wants to buy a Chinese-constrained unit (Figure 1.2a). He can only buy from a Chinese seller. One possible equilibrium is illustrated, where $\theta^C(u')$ is matched with $\phi^C(\pi^*)$. The Chinese buyer incurs a lower level of utility, u', because of the Chinese quotas since he has to pay a higher price, $\phi^C(z_1^*, \pi^*)$ (the lower price offered without the quotas, $\phi^{NC}(z_1^*, \pi^*)$, is not available to the Chinese buyer anymore because he cannot buy from the non-Chinese seller, ϕ^{NC}). In this example, the price the Chinese buyer has to pay, $P^C(z)$, to buy a unit in a block with 87% Chinese is a discrete jump upwards, compared to a unit in a block with $(87-\varepsilon)$ % Chinese. I discuss this discontinuity in prices below.

Figure 1.2b illustrates the case for the non-Chinese seller (ϕ^{NC}) in a Chinese-constrained unit. He can no longer sell to a Chinese buyer (the match between the grey bid and offer functions are not allowed). To sell his unit, one possibility is to sell to a non-Chinese buyer ($\phi^{NC}(\pi')$ matched to $\theta^{NC}(u^*)$) and incur a lower level of profit, π' (since he is paid a lower price then $\phi^{NC}(z_1^*, \pi^*)$). In this example, the price of units owned by non-Chinese sellers, $P^{NC}(z)$, jumps discretely downwards when the Chinese quota binds.

Figure 1.2c demonstrates the sense in which the ethnic housing quotas work like price discrimination. When the Chinese quota binds, the Chinese buyers face a negative supply shock because they cannot buy from non-Chinese sellers (P^C jumps up) while the non-Chinese sellers face a negative demand shock because they cannot sell to Chinese buyers (P^{NC} jumps down). When the Chinese quota binds, it is possible for prices to differ across ethnic groups because the quota prevents non-Chinese sellers from arbitraging price differences across ethnic groups (non-Chinese sellers cannot sell to Chinese buyers even though Chinese buyers are willing to pay a higher price than non-Chinese buyers).

Can we have discontinuities in prices at the quota?

A priori, one may not expect discontinuities in prices at the quota. If prices of a constrained unit jumped upwards at the quota (as in Figure 1.2a), would a Chinese buyer be willing to pay this premium to live just ε% above the quota? For units that are above the quota, consider the Chinese offer function, $\phi^C(\pi^*)$, in Figure 1.2a. If markets are truly competitive (there exists other Chinese sellers who own units with the characteristics $(z_1^*, ..., z_n^*)$), one of them could undercut the higher price, $\phi^C(z_1^*, \pi^*)$ and capture the entire market. For units right below the quota, if Chinese sellers knew that once the quota binds, there is a premium for their units, the probability of capturing this premium should already be priced into units that are ε% below the quota. In this case, prices should gradually increase as the percent of Chinese approaches 87% rather than a discrete jump upwards in prices at 87%.

Equivalently, if prices jumped downwards at the quota (as in Figure 1.2b) and non-Chinese sellers are forced to sell to non-Chinese buyers at a low price, $\phi^{NC}(z_1^*, \pi') = \theta^{NC}(z_1^*, u^*)$, would a non-Chinese seller be willing to accept a cut in prices? If markets were competitive (there exists other non-Chinese buyers who want units with characteristics $(z_1^*, ..., z_n^*)$, a non-Chinese buyer would outbid the price, $\theta^{NC}(z_1^*, u^*)$ and buy the unit. For units that are right below the quota, if non-Chinese buyers recognize that once the quota binds, there is a discrete downward jump in prices, this positive probability of the quota binding should be priced into units that are $\varepsilon\%$ below the quota. Hence, prices should gradually decrease as the percent of Chinese approaches 87%.

Based on the discussion above, there could be two reasons for discontinuities in prices at the quota: heterogenous goods and myopia. Consider Figure 1.2a again. In a market of heterogenous goods, the Chinese buyer who has chosen a unit is willing to pay a premium of $\theta^C(z_1^*, u') - \theta^C(z_1^*, u^*)$ because there are no unconstrained units that are exactly the same. Secondly, because of myopia, the Chinese who owns the unit right below the quota may not completely account for the probability his unit may become constrained and thus, enjoy a premium. This myopic behavior could explain why prices jump up instead of increasing gradually as the percent of Chinese approaches 87%. The discussion for Figure 1.2b is similar.

1.4 Data

I collected data on 35,718 actual transaction prices from the HDB website between April 2004 and November 2006. These are actual transaction prices, not imputed rents. From the same website, I downloaded monthly data on the quota status of apartment blocks, between March 2003 and October 2006. I match the quota status of the previous month to each transaction.[3]

I obtained data on the block level ethnic proportions (the key running variable in the regression discontinuity design), by hand matching more than 589,000 names to ethnicities using differences in the structure of Chinese, Malay and Indian names.[4] For example, most Chinese names only have 2 or 3 words and they also have very distinct last names (eg. Tan, Wong, Ng); Malay names are primarily Muslim names

[3]I repeated the analysis with a 3-month lag, instead of a 1-month lag and the main results are similar.

[4]In Singapore, a person's ethnicity and last name are determined by the father's ethnicity. This is recorded in the identity card and birth certificate.

since 100% of Malays in Singapore are Muslims (Singapore Census, 2000); Indian names are matched according to popular first and last names. In my experience, Indian and Malay names are the hardest to distinguish because there are 26% of Indians in Singapore who are Muslims and could adopt Arabic names, like Malays do (Singapore Census, 2000). As a check, my measure of ethnicity using names generates 78% "Chinese", 14% "Malays" and 8% "Indians", almost identical to the actual national proportions from the 2000 Census.

I have two datasets describing the distribution of housing by flat type. The first is a non-public dataset purchased from HDB which lists the number of each type of flat in all HDB blocks. The second dataset lists the type of flat sold (downloaded monthly from the HDB website on resale transactions).

Table 1.3 lists the summary statistics of the full dataset. There are 8,067 blocks and 35,718 resale transactions. The Chinese and Malay quotas bind for one-tenth of the sample and the Indian quotas bind for one-fifth of the sample.

The dataset is a combination of flow variables (monthly quota dummies and monthly transaction prices downloaded from the HDB website) as well as stock variables (ethnic proportions from the phonebook and types of flats in each block purchased from the HDB). I do not observe the ethnicity of the buyer and the seller which is important because they determine whether a transaction is allowed (for example, transactions of quota-constrained units involving buyers and sellers of the same ethnicity are always allowed).

1.5 Estimation

In this section, I analyze the effect of the quota on housing prices using a regression discontinuity approach (Angrist and Lavy, 1999, Hahn, Todd and Van der Klauw, 2001, Imbens and Lemieux, 2008). The regression discontinuity method relies on the step function of the quota status where units are unconstrained (the quota status is 0) below the upper bound on ethnic proportions and units are constrained (the quota status is 1) above the upper bound. The challenge in identifying the treatment effect is omitted variables. Even if the price effect of the Chinese quota was zero, the price of constrained units could be higher than the price of unconstrained units because areas with more Chinese amenities tend to attract more Chinese and are hence, more likely to be Chinese-constrained.

The regression discontinuity approach is to compare units right above and right below the quota. The treatment effect of the quotas is identified close enough to the discontinuity (the upper bound), assuming omitted variables are similar right above and right below the upper bound. For example, this assumes that the number of Chinese temples could be different above and below the upper bound, thereby generating price differences even absent the quota effect. But, the number of Chinese temples does not change discontinuously at the Chinese quota. If this assumption holds, comparing units right above and right below the quota offers one way to address the omitted variable problem. In practice, when the quotas started, the Housing Development Board (HDB) did not want to evict households from constrained areas. Hence, to this day, I still observe households above the quotas.

I estimate four sets of equations. Each is essentially regressing outcome variable(s) on a dummy for whether a quota is binding, controlling for smooth functions of ethnic proportions. I restrict my analysis to observations within 10% of the Chinese, Malay and Indian quotas respectively. The first three equations are at the month-block level and the final equation is at the block level. Recall that the dataset is comprised of flow and stock variables. All variables describing actual transctions (price, quality of units sold, quota dummies) are at the month-block level while all other variables (ethnic proportions, number of each flat type in a block) are stock variables. The first equation tests if there is in fact a step function in the probability of the quota binding.

$$\begin{aligned} QC_{bit} &= \alpha + \beta\, 1\,(percentC_{bi} \geq 0.87) + \sum_{k=1}^{4} \phi_k\,(percentC_{bi} - 0.87)^k \qquad (1.2) \\ &\quad + \sum_{k=1}^{4} \gamma_k 1\,(percentC_{bi} \geq 0.87) * (percentC_{bi} - 0.87)^k + \varepsilon_{bit} \end{aligned}$$

where QC_{bit} is a dummy for whether the (C)hinese quotas are binding for flats in block b, town i and month t (this is the assignment dummy obtained directly from the HDB data on quotas), $1\,(percentC_{bi} \geq 0.87)$ is a dummy for whether the percent of Chinese (data from the phonebook) is at or above the Chinese quota (87%), $(percentC_{bi} - 0.87)^k$ are kth order polynomials of the percent of Chinese, centered around the block quota.[5] The coefficient of interest is β, which represents the magnitude of the discontinuity at the quota.

The following 2 equations use the assignment dummy from HDB data (QC_{bit}) as

[5]The running variable is centred around the relevant quota (the running variable is *percent of Chinese in a block* $-$ 0.87) because I estimate two polynomials separately on the left and right of the quota. This ensures that the coefficient on the quota dummy represents the jump at the quota.

the key independent variable. Equation (1.3) only controls for smooth functions of the running variable, while equation (1.4) controls for other observable characteristics:

$$
\begin{aligned}
y_{bit} &= \alpha + \beta\, QC_{bit} + \sum_{k=1}^{4} \phi_k\, (percentC_{bi} - 0.87)^k && (1.3)\\
&+ \sum_{k=1}^{4} \gamma_k QC_{bit} * (percentC_{bi} - 0.87)^k + \varepsilon_{bit} \\
y_{bit} &= \alpha + \beta\, QC_{bit} + \sum_{k=1}^{4} \phi_k\, (percentC_{bi} - 0.87)^k && (1.4)\\
&+ \sum_{k=1}^{4} \gamma_k QC_{bi} * (percentC_{bi} - 0.87)^k + B_{bi}\, \delta + \tau_t + \omega_i + \varepsilon_{bit}
\end{aligned}
$$

where y_{bit} is the outcome variable for flats in block b, town i and month t; B represents other observable characteristics of the block (age of building, proportion of type 1 flats, type 2 flats etc.); τ_t and ω_i are month and town fixed effects. The standard errors in the second equation are clustered at the town level.

The fourth and final equation is aggregated to the block level:

$$
y_{bi} = \alpha + \beta\ percent\ QC_{bi} + \sum_{k=1}^{4} \phi_k percentC_{bi}^k + B_{bi}\ \delta + \omega_i + \varepsilon_{bi} \qquad (1.5)
$$

where $percent\ QC_{bi}$ is the proportion of months the Chinese quota is binding. For this analysis, I use observations within 10% of the quota but I do not control for smooth functions of polynomials separately on the right and left of the quota because I have aggregated the quota dummies (QC_{bit}) across months (my key regressor is not a dummy anymore).

I analyze three sets of outcomes. First, I test for differences in observable characteristics at the quota. Then, I test for discontinuities in the probability of the quota binding. The final set of regressions looks at the effect of the quota on proportion of units sold, price and quality of units sold.

I assume that the policy rule is perfectly enforced. Since these are public housing flats, all resale transactions need to be approved by the HDB. Part of the approval process involves checking that buyers and sellers of a transaction do not violate the ethnic quota rule.

1.6 Results

Figure 1.3 summarizes results from the estimation of equation (1.2) using observations within 10% of the quota. The dependent variable is a dummy for whether the quota is binding for block b in month t (this is actual quota data downloaded monthly from the HDB website). The main regressors are 4th order polynomials of ethnic proportions for block b and a dummy that is 1 when the ethnic proportion is above the quota (calculated using annual data from the phonebook). There is a discontinuity in the probability that each quota binds. The probability that the quota binds is positive even below the quota because of time series variation, a block can bind for a few months even though its annual ethnic proportion (a stock variable) is below the quota.

Table 1.4 reports results from a seemingly unrelated regression, with a system of outcome variables using equation (1.3). The outcome variables are*proportion type* 3, ..., *proportion type* 6 flats in each block.[6] Columns 1, 3 and 5 report results using all blocks that are within 10% of the Chinese, Malay and Indian quotas respectively. Columns 2, 4 and 6 report results using only blocks that were built before the policy started in 1989. The estimates show that the supply of flat types in Chinese- and Indian-constrained blocks are significantly different from unconstrained blocks but not so for blocks close to the Malay quota. These differences for Chinese and Indian blocks persisted since before the quota was introduced in 1989 because the estimates do not appear to be very different comparing columns 1 and 5 against columns 2 and 6. Chinese-constrained blocks have more type 3, type 4 and type 5 flats but fewer type 6 flats (all significant at 1%). Blocks close to the Malay quota are quite similar above and below the quota. Blocks close to the Indian quota have significantly more type 4 and type 6 flats, but significantly fewer type 3 flats. Whenever possible, I show estimates with and without controlling for these observable characteristics. Most findings are robust to the inclusion of these controls.

Tables 1.5 and 1.6 report results on price and quality. I use flat types as a measure of quality, where the dependent variable is an integer between 1 and 8 (there are 8 flat types). A higher number indicates a better flat type. Chinese-constrained units are 7% more expensive even though they are of a worse flat type. Malay-constrained units are 5% cheaper, perhaps because they tend to be of a worse flat type. Indian-constrained units are also cheaper (by 2%) even though the units sold tend to be of

[6] There are 8 types of flats. Higher types are more expensive. I only used 4 types in the regression because there are too few type 1, type 2, type 7 and type 8 flats in the resale market.

higher quality. Figures 1.4 and 1.5 show these results on price and quality.

Table 1.7 reports results of the quota impact on the proportion of units sold. An increase in the number of months that a quota binds significantly decreases the proportion of units sold. If the Chinese quota is binding for half a year more, the proportion of units sold will be lower by 3%. For Malays (Indians), the decrease is 5% (2%). The estimates for the Chinese and Malay quotas are significant at the 1% level. The estimates for the Indian quota is significant at the 10% level. Table 1.8 summarizes the impact of the quota on all 3 dimensions, quantity, price and quality.

1.7 Conclusion

Many desegregation policies take the form of quotas. This paper studies the impact of the ethnic housing quotas in Singapore that were designed to encourage residential desegregation amongst the three major ethnic groups, the Chinese, Malays and the Indians. I estimate the impact of the Chinese, Malay and Indian quotas on the price, quantity and quality of units sold. I find that all quotas significantly decrease the proportion of units sold. Malay-constrained units are 5% cheaper perhaps because the units sold are also of lower quality. The impact on the price and quality of Chinese- and Indian-constrained units are opposite. Chinese-constrained units are 7% more expensive even though the units sold are of significantly worse quality. Indian-constrained units are 2% cheaper even though the units sold are of a higher quality.

A priori, one may not expect quotas to cause discontinuities in characteristics of units sold at the cutoff. Finding these discontinuities in prices, quantity and quality, is consistent with a model with heterogenous goods and myopic agents. Unfortunately, without knowing the ethnicity of the buyer and the seller, nor the their preferences, it is hard to make normative judgements based on the estimated policy impact. Chapter 2 builds on these findings to estimate individual's taste for living with neighbors from the same ethnic group.

Bibliography

Angrist, Josh and Victor Lavy. 1999. "Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement." *Quarterly Journal of Economics*, 114(2): 533-75.

Commissie Blok. 2004. *28 689 Eindrapport Onderzoek Integratiebeleid*

Hahn, Jinyong, Petra Todd, and Wilbert van der Klaauw. 2008. "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design." *Econometrica,* 69: 201-209.

Housing Development Board (HDB). 2000. *Profile of Residents Living in HDB Flats.* Singapore: HDB.

Housing Development Board (HDB). 2006. *Annual Report.* Singapore: HDB.

Housing Development Board (HDB) website. http://www.hdb.gov.sg/

Imbens, Guido and Thomas Lemieux. 2008. "Regression Discontinuity Designs: A Guide to Practice." *Journal of Econometrics*, 127(2), 615-635.

Lum, Sau Kim and Tan Mui Tien. 2003. "Integrating Differentials: A Study on the HDB Ethnic Integration Policy." *Crestimes*, Singapore Center for Real Estate (CRES), 3(1): 5.

Singapore Department of Statistics. 2006. "Singapore Resident Population, 1990 - 2006."

Singapore Department of Statistics, Singapore 2000 Census. "Key Indicators of the Resident Population."

Table 1.1: Neighborhood and Block Level Quotas Relative to National Ethnic Proportions

	Neighborhood Quotas	Block Quotas	National Proportion (2000)
Chinese	84%	87%	77%
Malay	22%	25%	14%
Indian	10%	13%	8%

Source: 2000 Census (Singstat), Lum and Tan (2003)

Table 1.2: The Relationship between Quotas, Buyer Ethnicity and Seller Ethnicity

Binding Quota	Buyer Ethnicity	Seller Ethnicity	Status
Chinese	Chinese	Chinese	Allowed
	Non-Chinese	Non-Chinese	Allowed
	Non-Chinese	Chinese	Allowed
	Chinese	Non-Chinese	Not Allowed
Malay	Malay	Malay	Allowed
	Non-Malay	Non-Malay	Allowed
	Non-Malay	Malay	Allowed
	Malay	Non-Malay	Not Allowed
Indian	Indian	Indian	Allowed
	Non-Indian	Non-Indian	Allowed
	Non-Indian	Indian	Allowed
	Indian	Non-Indian	Not Allowed

Table 1.3: Summary Statistics

Variable	N	Mean	Std. Dev.	Level	Description
Price	35718	246477	70798	Month-Block	Average transaction price in a block (Singapore dollars)
Percent Sold	8067	4%	3%	Block	Percent of units in a block that was sold within the sample period
Chinese Quota	8067	11%	29%	Month-Block	Percent of units where Chinese quota binds
Malay Quota	8067	11%	28%	Month-Block	Percent of units where Chinese quota binds
Indian Quota	8067	21%	35%	Month-Block	Percent of units where Chinese quota binds
Percent Chinese	8067	78%	11%	Block	Percent of Chinese in a block
Percent Malay	8067	8%	6%	Block	Percent of Malay in a block
Percent Indian	8067	14%	9%	Block	Percent of Indian in a block
Age	35718	17.64	8.54	Block	Average age of HDB blocks
Percent Type 1	8067	0.05%	2%	Block	Percent of units in a block that is Type 1
Percent Type 2	8067	0.96%	8%	Block	Percent of units in a block that is Type 2
Percent Type 3	8067	23.28%	37%	Block	Percent of units in a block that is Type 3
Percent Type 4	8067	37.64%	34%	Block	Percent of units in a block that is Type 4
Percent Type 5	8067	25.10%	32%	Block	Percent of units in a block that is Type 5
Percent Type 6	8067	12.89%	32%	Block	Percent of units in a block that is Type 6
Percent Type 7	8067	0.01%	1%	Block	Percent of units in a block that is Type 7
Percent Type 8	8067	0.08%	3%	Block	Percent of units in a block that is Type 8

Table 1.4: Testing for Differences in Proportion of Flat Types, 10% Above and Below the Quota

Quota	Chinese		Malay		Indian	
Sample	Pre- and Post- Quota	Pre-Quota Only	Pre- and Post- Quota	Pre-Quota Only	Pre- and Post- Quota	Pre-Quota Only
Dependent variable	(1)	(2)	(3)	(4)	(5)	(6)
Proportion Type 3	0.01**	-0.04***	0.01	-0.02***	-0.02***	-0.05***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.004)	(0.01)
Proportion Type 4	0.01**	0.05***	-0.01	0.02***	0.02***	0.02***
	(0.005)	(0.01)	(0. .005)	(0.01)	(0.004)	(0.01)
Proportion Type 5	0.02***	0.03***	0.004	-0.00001	0.01**	0.03***
	(0.005)	(0.01)	(0.004)	(0.01)	(0.004)	(0.01)
Proportion Type 6	-0.05***	-0.05***	0.01	0.002	0.002	0.01**
	(0.004)	(0.01)	(0.005)	(0.01)	(0.004)	(0.005)
Observations	87512	64610	64963	50291	140200	121500

Notes: Results from a seemingly unrelated regression where the dependent variables are the proportion of flat types in a block. I control for two 4^{th} order polynomials, 2% to the left and right of the quota respectively. Only the coefficient on the quota dummy is reported. Columns 1, 3, and 5 use all blocks. Columns 2, 4 and 6 use only blocks built before the quota was implemented.
Standard errors in parentheses.
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 1.5: Impact on Price, 10% Above and Below the Quota

Quota	Chinese		Malay		Indian	
Dependent variable	ln price	ln price	ln price	ln price	ln price	ln price
	(1)	(2)	(3)	(4)	(5)	(6)
Quota dummy	0.07***	0.04***	-0.05***	-0.03***	-0.02**	-0.03**
	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)
Ethnic proportion	-0.49***	-0.09	0.02	-0.19*	-1.87***	-0.37**
	(0.16)	(0.12)	(0.18)	(0.10)	(0.17)	(0.15)
(Ethnic proportion)^2	9.00***	-0.25	-9.20***	0.09	-24.31***	0.27
	(3.47)	(0.28)	(3.50)	(0.18)	(3.00)	(0.17)
(Ethnic proportion)^3	198.32***	-1.59	-17.1	-5.20**	163.33***	-2.69
	(32.68)	(2.40)	(32.60)	(2.36)	(33.92)	(1.92)
(Ethnic proportion)^4	354.87	11.08*	858.17*	7.15	2993.26***	2.58
	(467.81)	(5.49)	(447.36)	(6.99)	(428.39)	(3.63)
Quota*Ethnic proportion	-0.60*	6.05	0.59**	-7.53	1.20***	44.17*
	(0.32)	(25.72)	(0.28)	(17.27)	(0.25)	(23.01)
Quota*(Ethnic proportion)^2	25.40***	54.11*	6.47	-3.13	15.26***	-15.35
	(8.50)	(28.80)	(6.82)	(18.96)	(5.82)	(26.58)
Quota*(Ethnic proportion)^3	-211.53***	213.61	-16.5	500.67	-43.37	294.73
	(64.11)	(327.28)	(51.24)	(296.45)	(48.07)	(229.23)
Quota*(Ethnic proportion)^4	-3971.24***	-816.81	512.29	-548.76	-1865.46**	-294.25
	(1150.87)	(604.56)	(868.09)	(848.86)	(769.96)	(486.67)
Block level controls	No	Yes	No	Yes	No	Yes
Town fixed effects	No	Yes	No	Yes	No	Yes
Month fixed effects	No	Yes	No	Yes	No	Yes
Observations	19314	19314	14862	14862	32114	32114
R-squared	0.02	0.8	0.01	0.75	0.01	0.78

Notes: Columns 2, 4 and 6 control for block level characteristics (age, age-squared, proportions of flat types), month and town fixed effects. Standard errors for columns 2, 4 and 6 are clustered at the town level.
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 1.6: Impact on Quality of Unit Sold in a Block, 10% Above and Below the Quota

Quota	Chinese	Malay	Indian
Dependent variable	Flat Type Sold	Flat Type Sold	Flat Type Sold
	(1)	(2)	(3)
Quota dummy	-0.05*	-0.08***	0.04*
	(0.03)	(0.03)	(0.02)
Ethnic proportion	-2.51***	0.83	-4.42***
	(0.48)	(0.60)	(0.52)
(Ethnic proportion)^2	33.77***	1.51	-63.29***
	(10.41)	(11.61)	(9.30)
(Ethnic proportion)^3	653.33***	-106.64	461.34***
	(98.02)	(108.25)	(105.28)
(Ethnic proportion)^4	1273.14	-199.71	8818.55***
	(1403.38)	(1485.44)	(1329.68)
Quota*Ethnic proportion	0.92	2.61***	4.10***
	(0.97)	(0.92)	(0.77)
Quota*(Ethnic proportion)^2	26.72	15.85	60.28***
	(25.49)	(22.66)	(18.06)
Quota*(Ethnic proportion)^3	-970.03***	-244.69	-331.34**
	(192.34)	(170.13)	(149.21)
Quota*(Ethnic proportion)^4	-9300.65***	975.02	-7780.61***
	(3452.52)	(2882.48)	(2389.90)
Observations	19314	14862	32114
R-squared	0.01	0.003	0.004

Standard errors in parentheses.
* significant at 10%; ** significant at 5%; *** significant at 1%

Notes: The dependent variable is an integer between 1 and 8. A higher number corresponds to higher quality.

Table 1.7: Impact on Quantity of Units Sold in a Block, 10% Above and Below the Quota

Quota	Chinese	Malay	Indian
Dependent variable	Proportion Sold	Proportion Sold	Proportion Sold
Quota dummy	-0.005***	-0.009***	-0.004*
	(0.002)	(0.001)	(0.002)
Ethnic proportion	-3.914	-5.009	0.35
	(6.474)	(4.793)	(0.389)
(Ethnic proportion)^2	3.306	33.964	-3.049
	(5.651)	(32.276)	(5.857)
(Ethnic proportion)^3	dropped	-98.679	9.224
		(94.512)	(35.989)
(Ethnic proportion)^4	-0.705	104.934	2.87
	(1.267)	(101.604)	(76.785)
Block level controls	Yes	Yes	Yes
Town fixed effects	Yes	Yes	Yes
Month fixed effects	Yes	Yes	Yes
Observations	3948	2896	6322
R-squared	0.1	0.12	0.11

Notes: Block -level data. The dependent variable is the proportion of units sold in a block. Block level controls include age, age-squared and proportions of flat types.
Standard errors are clustered at the town level.
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 1.8: Summarizing the Impact of the Quota on Quantity, Price and Quality

	Chinese Quota	Malay Quota	Indian Quota
Quantity	Lower	Lower	Lower
Price	Higher	Lower	Lower
Quality	Lower	Lower	Higher

Figure 1.1: A Market Equilibrium Without Quotas

Figure 1.1a: A Consumer Equilibrium

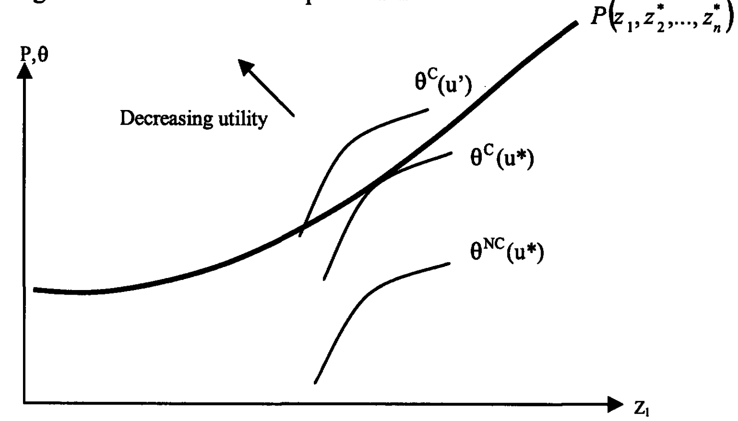


Figure 1.1b: A Producer Equilibrium

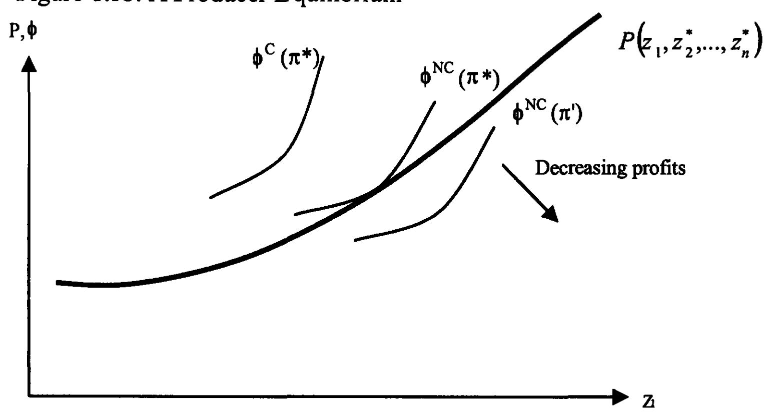
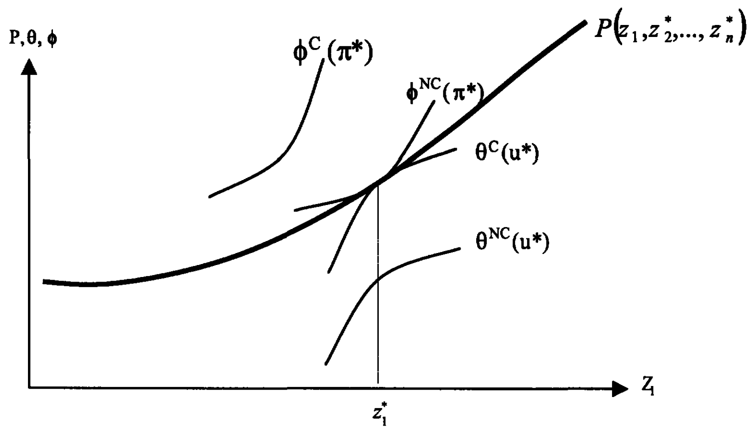


Figure 1.1c: A Market Equilibrium



Notes: These three panels illustrate one possible market equilibrium for a product with n characteristics ($z_1,\ldots,z_n$). Figures 1.1a and 1.1b plot "indifference surfaces" and "isoprofit surfaces" projected onto the θ,ϕ–z_1 plane cut at optimal values of other characteristics $(z_2^*,\ldots, z_n^*)$. The superscripts C and NC denote Chinese and non-Chinese. θ(u) is a buyer's bid function. (the expenditure he is willing to pay for different levels of z_1 to achieve utility u); $\phi(\pi)$ is a seller's offer function (the price he is willing to accept to achieve profit, π); p(z) is the hedonic price function (the equilibrium market price). Each point on the hedonic price function is a tangency between a buyer's bid function and a seller's offer function. Here, a non-Chinese seller is matched to a Chinese buyer in equilibrium.

Figure 1.2: The Impact of Chinese Quotas on Prices of Constrained Units

Figure 1.2a: The Impact on Chinese Buyers, θ^C

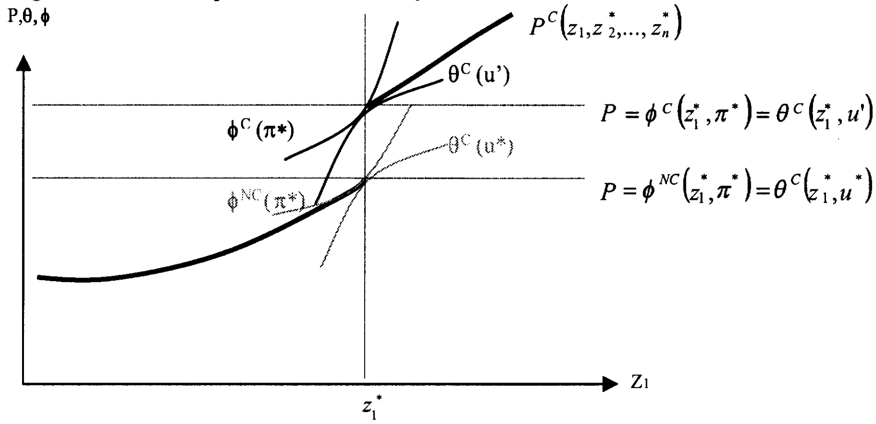


Figure 1.2b: The Impact on Non-Chinese Sellers, ϕ^{NC}

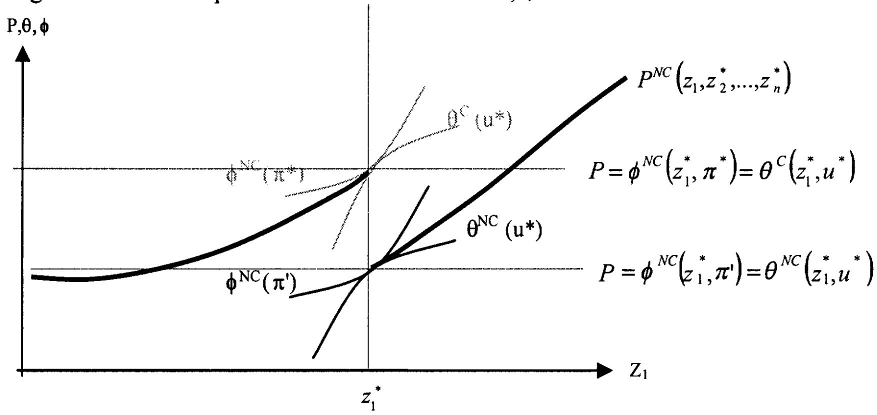
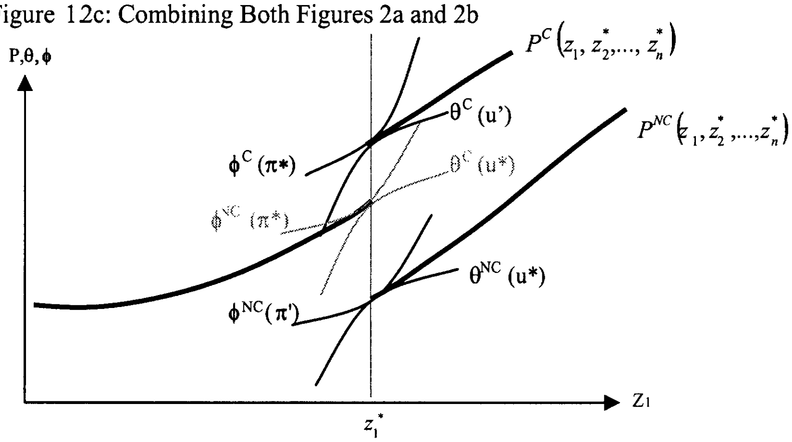
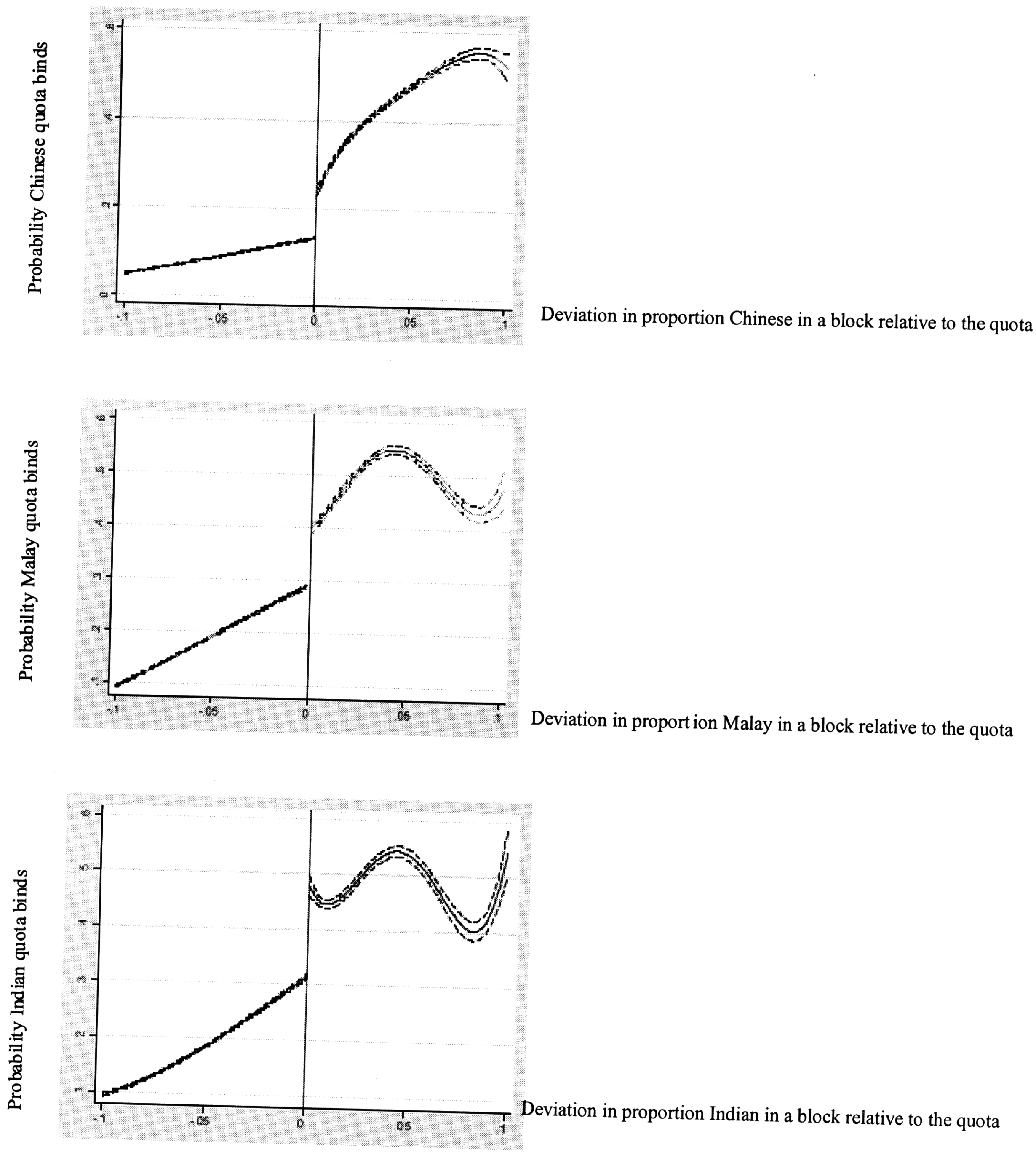


Figure 1.2c: Combining Both Figures 2a and 2b



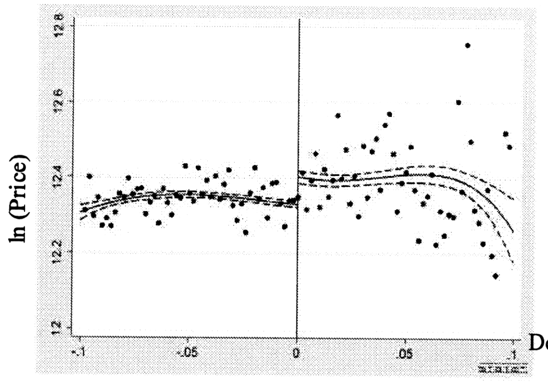
Notes: These three panels illustrate one possible consequence of the Chinese block quotas for units that are just above the quota. The horizontal axis, z_1, represents percent of Chinese in a block and z_1^* represents the Chinese block quota (87%). Units to the right of the vertical line are constrained. For Chinese-constrained units, non-Chinese sellers cannot sell to Chinese buyers (the grey offer function, $\phi^{NC}(\pi^*)$ cannot be matched to the grey bid function, $\theta^C(u^*)$). In this example, the Chinese buyer has to pay a higher price (Figure 1.2a) and the non-Chinese Seller has to accept a lower price (Figure 1.2b). Figure 1.2c combines both figures. Above the Chinese quota, prices are allowed to differ across groups (price discrimination).

Figure 1.3: Testing for Discontinuity in the Probability that the Quota Binds, 10% Above and Below the Quota

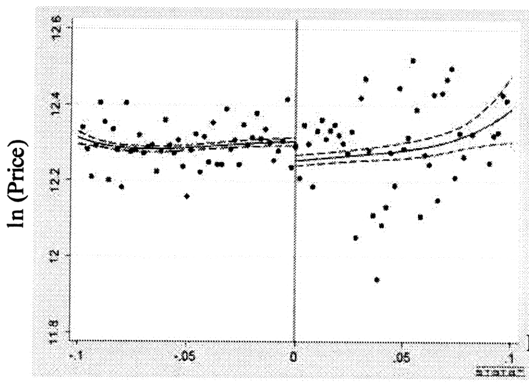


Notes: Eachpanel in this figure is constructed using the following procedure for observations within 10% of the ethnic quotas: (i) regress Q (a dummy for whether the quota is binding) on smooth functions of the corresponding running variable (4th order polynomials), separately, once to the left and once to the right of the quota; (ii) plot the predicted probabilities above and below the quota separately. Repeat the exercise for the Malay quotas and Indian quotas. The dashed lines represent 95% confidence intervals.

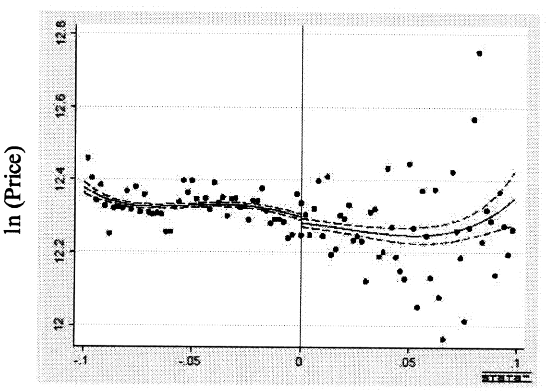
Figure 1.4: Impact on lnPrice, 10 % Above and Below the Quota



Deviation in proportion Chinese in a block relative to the quota



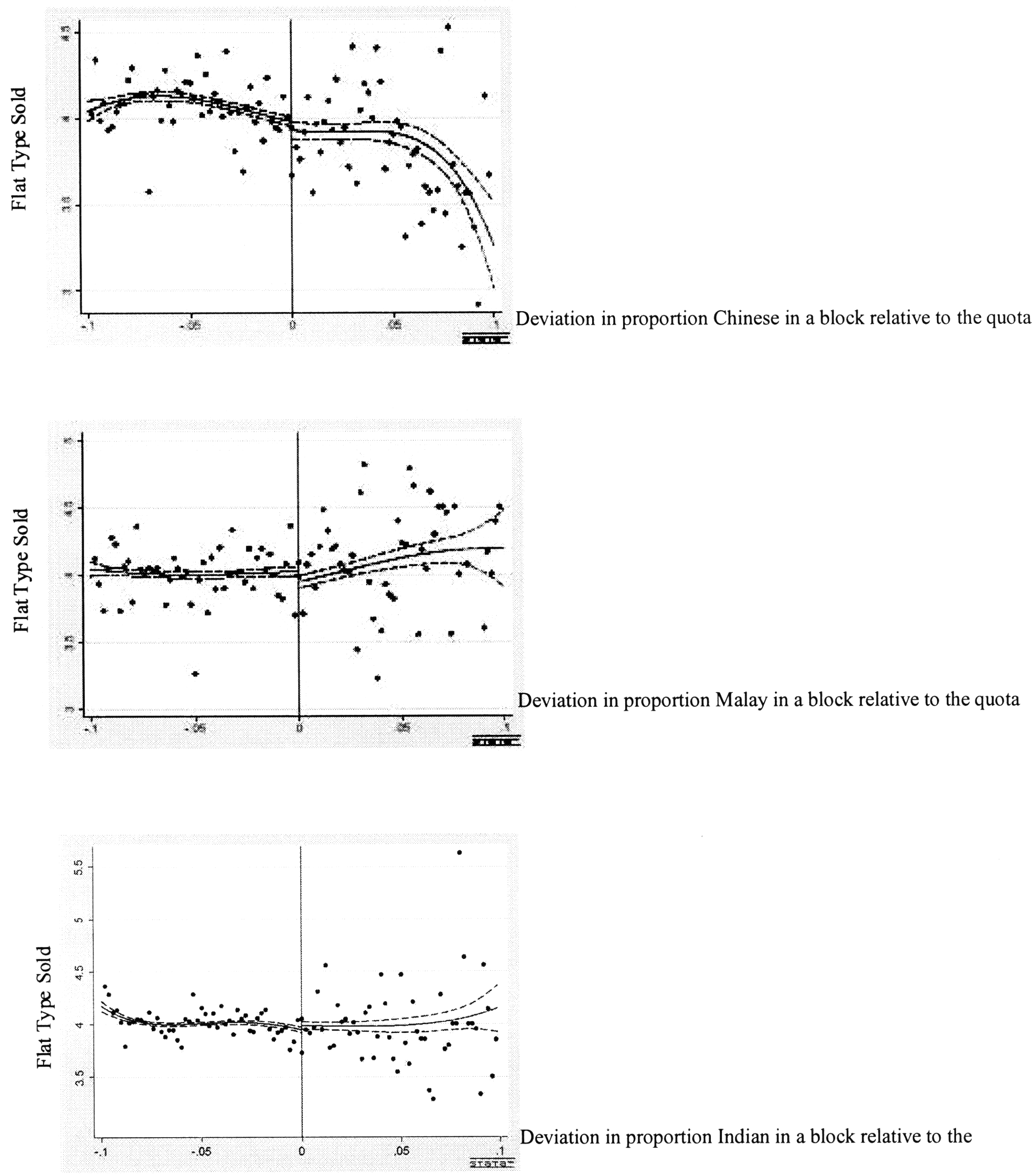
Deviation in proportion Malay in a block relative to the quota



Deviation in proportion Indian in a block relative to the quota

Notes: Each panel in this figure is constructed using the following procedure for observations within 10% of the ethnic quotas: (i) regress *ln(Price)* on smooth functions of the corresponding running variable (two separate 4th order polynomials, one to the left and one to the right of the quota) and a dummy that is one when the corresponding block quota is binding (ii) plot the predicted *ln(Price)* above and below the quota separately (iii) plot means of *ln(Price)* for each 0.2% bin. Repeat the exercise for the Malay quotas and Indian quotas. The dashed lines represent 95% confidence intervals.

Figure 1.5: Impact on Type of Flat Sold, 10% Above and Below the Quota



Notes: The dependent variable is an integer between 1 and 8, describing the type of flat sold. A higher flat type is better. Each panel in this figure is constructed using the following procedure for observations within 10% of the ethnic quotas: (i) regress *Flat Type Sold* on smooth functions of the corresponding running variable (two separate 4th order polynomials, one to the left and one to the right of the quota) and a dummy that is one when the corresponding block quota is binding; (ii) plot the predicted *Flat Type Sold* above and below the quota separately (iii) plot means of for each 0.2% bin. Repeat the exercise for the Malay quotas and Indian quotas. The dashed lines represent 95% confidence intervals.

Chapter 2

Estimating Ingroup Preferences Using Ethnic Housing Quotas in Singapore

2.1 Introduction

Many ethnically diverse countries try to promote desegregation in areas such as education, employment, immigration and housing. For example, in Chicago, the Gautreaux program (the predecessor of the Moving Towards Opportunity program) tried to encourage residential desegregation using housing vouchers (Jennan, 2000). In Singapore, the government instituted a set of controversial ethnic housing quotas to encourage residential integration. In education, immigration and employment, similar integration policies are constantly being debated. Economists have studied extensively the consequences of segregation (Cutler and Glaeser, 1997; Ananat, 2007). In this paper, I explore one potential cause of residential segregation–ingroup preferences.[1] In this paper, I exploit variation from a natural experiment in Singapore in a structural model to estimate taste for living with one's own ethnic group members (ingroup preferences).

Understanding ingroup preferences is important for at least three reasons. First, residential segregation is a prominent feature in many ethnically diverse countries but we know relatively less about its causes. Second, we need to understand whether residential segregation is driven by tastes for ethnic interactions or tastes for ethnic-specific amenities because the policy implications are different: If segregation is driven by tastes for ethnic ingroups, integration policies should focus on incentivizing resi-

[1] Another cause of residential segregation is discrimination (Cutler, Glaeser and Vigdor, 1999). I discuss the implications of discrimination below.

dential location choices. If, instead, segregation patterns arise because ethnic group members share the same tastes for amenities (correlated tastes), then reducing the spatial clustering of ethnic-specific amenities can encourage integration. Finally, my approach allows me to recover the elasticities of demand for ingroup interactions. These are crucial inputs for measuring potential deadweight losses from regulating location choices, which will have to be weighed against potential benefits of integration.

To estimate ingroup preferences over residential locations, I build and estimate a discrete choice model à la Berry, Levinsohn, and Pakes (1995) in which individuals choose residential neighborhoods as a function of the proportion of ethnic ingroup members in a neighborhood. The empirical challenge is that the explanatory variable of interest, neighborhood ethnic proportions, could be correlated with unobserved ethnic-specific neighborhood quality (as measured by ethnic-specific amenities, for example). An amenity such as mosques could be more attractive to a specific ethnic group if a majority of that ethnic group is Muslim. In this case, just addressing omitted variables that are common across groups, using a neighborhood fixed effect, for example, is insufficient if some neighborhoods have systematically more mosques that will attract a specific ethnic group. One way to address ethnic-specific omitted variables is to use some within neighborhood, across group variation, such as a model with neighborhood-by-group fixed effects.

The ethnic housing quotas in Singapore provide a source of identification. They were introduced in 1989 to encourage residential integration amongst the three major ethnic groups in Singapore – Chinese (77%), Malays (14%) and Indians (8%) (Singapore Census, 2000). The policy is a set of upper bounds on block level and neighborhood level ethnic proportions. In practice, the Housing Development Board (HDB) did not want to evict owners in units that were in violation of the quotas. To this day, there exist units above the upper bound. Any transactions that forced these blocks and neighborhoods farther above the upper bound, however, would be barred. For example, when Chinese quotas are binding, non-Chinese sellers cannot sell to Chinese buyers because this transaction increases the Chinese proportion farther above the Chinese quota. These ethnic-specific restrictions on transactions play a central role in my analysis.

The quotas help identification of ingroup preferences in two ways. First, they imply that prices faced by different ethnic groups differ in equilibrium when the quota binds. The intuition is similar to that of price discrimination models. In such models, different groups can be charged different prices in equilibrium as long as there are no arbitrage opportunities. The quotas impose ethnic-specific restrictions on transactions that prevent some sellers from arbitraging away price differences across ethnic groups (when the Chinese quotas bind, non-Chinese sellers cannot sell to Chinese buyers). Therefore, prices can differ across ethnic groups within the same neighborhood when the quotas bind. Since prices are positively correlated with quality, I exploit information from group-specific prices to recover group-specific neighborhood quality.

The second reason the ethnic quotas help identification of ingroup preferences is that the policy rule is a step function of ethnic proportions. Units above the upper bound on ethnic proportions are constrained (the quota status is 1) while units below the upper bound are unconstrained (the quota status is 0). This step function is an ideal set up for a regression discontinuity design (Angrist and Lavy, 1999). The idea behind regression discontinuity is that close enough to the upper bound, omitted variables are assumed to not change discontinuously at the upper bound.

I use this step function to construct an instrument for group-specific prices. While the quota status is correlated with prices, it could also be correlated with neighborhood quality, thus, violating the exclusion restriction. [2] To construct an instrument for group prices, I first estimate ethnic proportions using a set of exogenous instruments.[3] Then, I assign a quota status of 1 if the predicted ethnic proportions are above the upper bound as defined by the policy rule, and 0 otherwise. Constructing instruments using this step function is akin to the regression discontinuity identification strategy: Exogenous characteristics used to estimate the quotas could affect prices but the effect will not be discontinuous at the upper bound of the quota.

I embed this regression discontinuity identification strategy in a structural demand estimation framework à la Berry, Levinsohn and Pakes (1995), hereafter BLP. In practice, I estimate three BLP models, one for each ethnic group, using the method of simulated moments. Because the quotas allow price discrimination to exist in this market, product attributes can vary across groups. More importantly, I allow unobserved neighborhood quality to vary by ethnic group.

To implement this analysis, I collected individual data on residential location choices by matching names from the 2005 and 2006 Singapore residential phonebooks. I hand matched more than 589,000 names to ethnicities.[4] I also collected data on

[2] Chinese quotas are more likely to bind in neighborhoods with a high Chinese quality.

[3] I use exogenous characteristics of nearby neighborhoods as instruments. I also use historical data on early ethnic settlements to instrument for ethnic proportions.

[4] In Singapore, one's ethnicity depends on the father's ethnicity. Also, inter-ethnic marriages are fairly rare (10%) mostly because the different ethnic groups have different religions. The Chinese

neighborhood characteristics, prices and quotas for 170 neighborhoods. My price data consists of weighted averages of group-specific prices.[5]

I find that all groups have strong preferences for living with at least some other members of their ethnic group. Moreover, the Chinese and the Indians have ingroup preferences that are inverted U-shaped. For example, the average Indian household living in a neighborhood with 5% Indians (the 10th percentile) is willing to substitute to a neighborhood that is 2.56km *further* from the closest subway station in exchange for living in a neighborhood with a 1 standard deviation (3%) increase in the proportion of Indians. However, after a neighborhood reaches 8% Indians and 43% Chinese respectively, Indians and Chinese would rather add a new neighbor from the other groups. These estimates are significant at the 5% level. Malays appear to have strong ingroup preferences, although the estimates for their taste parameters are not significant.

This paper makes two contributions. First, the estimates on ingroup preferences have implications for research and policy. The finding that some households in Singapore exhibit ingroup preferences that are inverted U-shaped suggests that models that estimate ingroup preferences assuming monotonicity could be biased. In addition, policy makers could use these estimates on ingroup preferences to determine subsidies for housing voucher programs that aim to encourage desegregation.

The second contribution of this paper is that I combine policy induced variation from a natural experiment with structural methods. I embed the regression discontinuity identification strategy within a structural demand estimation framework by using the step function to construct exogenous quota variables as instruments for group-specific prices.

The backbone of the empirical analysis rests on the literature on discrete choice theory (see for example McFadden, 1974; Berry, Levinsohn, and Pakes, 1995, 2004). This is not the first paper that estimates multiple unobserved characteristics (see Das, Olley, and Pakes, 1993; Athey and Imbens, 2007). It demonstrates that in markets with price discrimination, variation in prices across groups can provide additional variation for identification.[6]

Other contributors to this paper include the literature on social interactions and

---

are primarily Buddhists (54%) and Christians (17%), the Malays are 100% Muslims and the Indians are primarily Hindus (55%) and Muslims (12%). (Singapore Census, 2000)

[5] Not observing actual group-specific prices is unfortunate, but this is a limitation that is specific to my dataset. For the rest of the paper, I discuss the identification strategy assuming I observe group-specific prices. In the data section, I discuss how I estimate group-specific prices using a weighted average of prices and observed weights.

[6] This can be useful if the product attributes studied lack variation across markets.

residential segregation that dates back to Manski (1993), Schelling (1971) and more recently, Cutler, Glaeser and Vigdor (1999) and Card, Mas and Rothstein (2007). Bayer, Ferreira and McMillan (2007) and Bajari and Kahn (2005) estimate ethnic preferences using random coefficients. Both papers estimate a rich set of taste parameters using many demographic variables. My paper complements these two papers by identifying policy variation that I exploit to allow neighborhood quality to vary across groups. This accounts for ethnic-specific omitted variables thereby identifying ingroup preferences from correlated tastes for ethnic-specific neighborhood quality.

In the next section, I discuss the background of ethnic quotas in Singapore. Then, I describe the data (Section 3) and the results from the regression discontinuity analysis (Section 4). I then build a model of individual utility over residential locations with ethnic-specific prices (Section 5), discuss estimation of the structural model (Section 6) and present the results (Section 7). Finally, I conclude in Section 8.

2.2 Background

Singapore is a multi-ethnic country with a population of 4.5 million (Singapore Department of Statistics, 2006). The three major ethnic groups are the Chinese (77%), the Malays (14%) and the Indians (8%). The Chinese have the highest median monthly income (S\$2335), followed by the Indians (S\$2167) and the Malays (S\$1790) (Singapore Census, 2000).

Public housing is the most popular choice of housing in Singapore with 82% of the resident population living in public housing (HDB, 2006). The units are built and managed by the Housing Development Board (HDB). There are three ways Singapore residents can live in a HDB flat. They may apply through the primary allocation system for new HDB flats, they may purchase existing HDB flats in the resale market or they may rent. The rental market is negligible: 98% percent of the HDB flats are owner-occupied (HDB, 2006). This paper focuses on the resale market which is the relevant market for the ethnic quotas. Relative to the primary market which is heavily regulated, the resale market functions as an open market.

To understand the ethnic quotas, it is important to understand the geography of housing markets in Singapore. The smallest spatial unit is an HDB *flat*. A group of HDB flats constitute an HDB *block*. A group of HDB blocks make up a *neighborhood*. An HDB block is comparable to a US Census block group, with an average of 70 households. An HDB neighborhood is comparable to a US Census tract, comprising an average of 60 HDB blocks. Throughout my analysis, I define a *market* as a cluster of neighborhoods.

The government of Singapore introduced the Ethnic Integration Policy to address the "problem" of the increase in the "concentrations of racial groups" in HDB estates (Parliamentary debate, 1989). The policy was announced in a parliamentary debate on February 16, 1989 and implemented starting March 1, 1989. The Policy is a set of quotas at the block and neighborhood level. Table 2.1 lists the quotas, in comparison to the 2000 national ethnic proportions. Neighborhood quotas are 2% to 8% above national ethnic proportions. Block quotas are 3% above neighborhood quotas, allowing more flexibility at the block level (blocks can be more segregated than neighborhoods). In practice, the HDB did not want to evict owners in existing units that were in violation of the quotas. To this day, there exist blocks and neighborhoods above the quota.

The quotas are upper bounds on ethnic proportions to *prevent* HDB communities that are already segregated from becoming more segregated. Once a community hits the upper bound, transactions that make the community more segregated will be blocked. However, other transactions will be allowed. In particular, transactions involving buyers and sellers from the same ethnicity will always be allowed because this does not increase the ethnic proportion. As an example, Table 2.1 shows that the Chinese neighborhood quota is set at 84%. Once Chinese make up more than 84% of the neighborhood population, Chinese buyers can no longer buy from non-Chinese sellers because this increases the proportion of Chinese in that neighborhood. Table 2.2 lists the types of transactions barred by each ethnic quota. The important thing to note is that once a Chinese quota binds, the Chinese buyers can no longer buy from non-Chinese sellers. Similarly for Malay and Indian quotas. This group-specific restriction prevents arbitrage and thus allows prices to differ across groups in equilibrium.

2.3 Data

I use data covering 170 neighborhoods and 7 markets for resale transactions in the public housing market in Singapore, between April 2005 and March 2006. This dataset encompasses virtually all of Singapore.[7] A neighborhood is comparable to a US Census Tract (4500 households, on average). A market is a cluster of neighbor-

[7] The analysis only focuses on the public housing market which represents 82% of the residents in Singapore. To the extent that households with strong ingroup preferences have sorted away from being regulated by the quotas and into the private housing market, the estimates of ingroup preferences from the resale market would be a lower bound.

hoods, categorized according to the Straits Times Real Estate Classifieds (the leading English newspaper in Singapore). The number of neighborhoods in each market varies from 12 to 38, with a mean of 24 and a standard deviation of 10.

Prices

I collected data on 25,182 actual transaction prices from the HDB website between March 2005 and April 2006. I aggregate these monthly transaction level prices into average, annual neighborhood prices. These are actual purchase prices, not rents. Unfortunately, I do not observe the ethnicity of the buyers and sellers. Hence, for each neighborhood j, I only observe a weighted average of the actual group-specific prices. I discuss how I estimate group prices using observed weights in the following sub-section.

Neighborhood characteristics

Neighborhood characteristics include ethnic proportions (of the initial *stock* of residents in a neighborhood), school quality, access to public transportation, the average age of HDB buildings and the average number of rooms in HDB buildings. To calculate ethnic proportions, I hand matched more than 589,000 names to ethnicities using differences in the structure of Chinese, Malay and Indian names. For example, most Chinese names only have 2 or 3 words; Malay names are primarily Muslim names since 99% of Malays in Singapore are Muslims (Singapore Census, 2000); Indian names are matched according to popular first and last names. I collected the remaining neighborhood attributes from online street directories, the HDB website and a non-public dataset purchased from HDB. See the Data Appendix for definitions of these variables and their sources.

Choice data

I collected data on individual residential location choices by matching names from the 2005 and 2006 Singapore residential phonebooks.[8] I define *movers*, as individuals whose postal code in 2005 did not match with their postal code in 2006. A postal code uniquely identifies an HDB block. In cases with multiple names, I match movers to their choices randomly.[9] There are 16,092 movers.[10] Using the ethnicity of movers,

[8] In principle, one could use just 1 phonebook to analyze the residential location choices of all residents in Singapore (using the choice data of both movers and stayers). I focus on using the choices of movers instead of all residents because choices of movers are actual location choices involving costly transactions.

[9] Multiple names will not affect my analysis because only the ethnicity of the buyer matters in the calculation of the ethnic shares. For my purposes, ethnicity depends only on names.

[10] The actual number of transactions (25,182) is higher than the number of movers (16,092). This could be due to spelling errors between names in the 2005 phonebook and the 2006 phonebook. I assume that the spelling errors are random such that the choice data of movers is not a selected sample.

I calculated ethnic shares. For example, the percent of the *flow* of Chinese buyers who chose a neighborhood. There are 13 neighborhoods with no movers in my sample period at all, 2 neighborhoods with no Chinese movers, 4 with no Malay movers, 6 with no Indian movers and 1 with no Malay nor Indian movers. For these neighborhoods, I assign their shares to be the minimum share for each ethnic group.[11] Note that ethnic proportions describe the ethnic distribution of the *stock* of residents while ethnic shares refer to the *flow* of movers. In my analysis, I use *ethnic shares* as a proxy for demand and *ethnic proportions* as the primary explanatory variable. The assumption is that the flow of movers is so small that the ethnic proportion of the stock of residents is essentially constant within a year.

Early ethnic settlements

I use data on early 19th century ethnic settlements in Singapore to instrument for ethnic proportions. The Jackson Plan, which was commissioned in the 1820s, specified that the west of the Singapore River should be reserved for Chinese and Indian communities while the east of the river was reserved for the Malay communities (see Figure 2.1a). I argue that the historical assignment of ethnic settlements to opposite banks of the river is an important instrument for current ethnic proportions.

Figure 2.1a shows the map of early 19th century Singapore according to the Jackson Plan (Crawfurd, 1828). The Jackson Plan was formulated by a committee led by Lieutenant Philip Jackson to set up ethnic functional subdivisions within the growing port city. The map illustrates early settlements in Singapore around the Singapore River. Four separate residential areas were designated for the Chinese, Malays, Indians and Europeans. The Malay and European towns were to the east of the Singapore River while the Chinese and Indian areas were to the west of the river. Figure 2.1b shows the distribution of existing quota-constrained neighborhoods. Malay-constrained neighborhoods are primarily in the east while the Chinese neighborhoods are in Central, South and West Singapore and the Indian neighborhoods are in Central and North Singapore. This suggests that the Malay neighborhoods expanded to the east of the river, perhaps because early Malay ethnic settlements were already there. I define a dummy variable that is 1 when the entire area of the neighborhood is in the east of the eastern end of early Malay settlements, and 0 otherwise.

Table 2.3 lists the summary statistics of the full dataset. There are 170 neighborhoods. The ethnic shares are very low (the means for all groups are below 0.5%)

[11]Because the estimation involves the inversion of ethnic shares, shares of neighborhoods that are zero are not invertible. As an approximation, I assign minimum shares to these neighborhoods.

indicating that the flow of buyers is very low. The Chinese quotas bind for almost one-fifth of the sample, the Malay quotas bind for one-tenth of the sample and the Indian quotas bind for a quarter of the sample.

2.3.1 Estimated Prices

I estimate ethnic-specific prices for each neighborhood using a weighted average of prices at the block level as well as the observed ethnic weights (from the data on movers from the phonebook). This estimation procedure essentially solves a system of equations where the variables of interest are the block level ethnic weights and the unknowns are the neighborhood level group prices. The assumption is that within a neighborhood, controlling for block level characteristics, prices vary across blocks only because the ethnic weights vary (a block with buyers who are 20% Chinese and 80% Malay will have a different average price from a similar block with 80% Chinese buyers and 20% Malay buyers).

I estimate the following equation for the average price of block b in neighborhood j:

$$\ln \bar{P}_{bj} = \pi^C I_j * w_{bj}^C + \pi^M I_j * w_{bj}^M + B_{bj} * I_j \gamma + \pi I_j + v_{bj} \tag{2.1}$$

where w_{bj}^C and w_{bj}^M are the Chinese and Malay buyer weights, B_{bj} is a set of block-level characteristics (the block quotas, the number of 1-room flats, 2-room flats etc.), I_j is a neighborhood dummy and the Indians are the omitted group. Notice that the neighborhood dummy is interacted with each explanatory variable. Essentially, for each neighborhood, I am solving a system of N_j equations and 3 unknowns for each neighborhood j where N_j is the number of blocks in neighborhood j and the 3 unknowns are the Chinese, Malay and Indian prices for neighborhood j. Using estimates from equation (2.1), I substitute $w_{bj}^C = 1$, $w_{bj}^M = 0$ to obtain the Chinese price, $\hat{P}_j^C$ and likewise for the Malay and Indian prices.[12]

2.4 Regression Discontinuity

In this section, I analyze the effect of the quota on observed average prices and estimated group prices using a regression discontinuity approach (Angrist and Lavy,

[12] I calculate $\hat{P}_j^C = \exp\left(\ln \overline{P}\left(\hat{\pi}, \hat{\gamma}; w_j^C = 1\right)\right)$. I use log of prices to ensure that the price estimates are positive. In ongoing work, I estimate group prices using average prices (instead of log prices) but constraining price estimates to be positive.

1999). First, I discuss the regression discontinuity setup in the context of the quotas. Then, I give a simple conceptual framework to analyze the effect of the quotas on ethnic-specific prices and average prices. Finally, I discuss the regression discontinuity results of the quota effect on average and estimated ethnic prices. A key identification assumption in my paper is that the quotas generate ethnic-specific prices by preventing arbitrage and these ethnic prices have information on unobserved ethnic-specific quality. I argue that even though I do not observe prices by ethnicity in my data, findings from the regression discontinuity analysis using average prices is suggestive that there is no arbitrage.

2.4.1 Regression discontinuity analysis

The regression discontinuity method relies on the step function of the quota status where units are unconstrained (the quota status is 0) below the upper bound on ethnic proportions and units are constrained (the quota status is 1) above the upper bound. The challenge in identifying the treatment effect is omitted variables. Even if the quota effect on prices was zero, the price of all units above the quota could be higher than the price of all units below the quota because neighborhoods where the quota binds tend to have higher neighborhood quality.[13] The regression discontinuity approach is to compare units right above and right below the quota. The treatment effect of the quotas is identified close enough to the discontinuity (the upper bound), assuming omitted variables are similar right above and right below the upper bound. For example, this assumes that the number of mosques (a measure of unobserved Malay neighborhood quality) could be different above and below the upper bound, thereby generating price differences even absent the quota effect. But, the number of mosques does not change discontinuously at the upper bound of the Malay neighborhood quota. If this assumption holds, comparing units right above and right below the quota offers one way to address the omitted variable problem. In practice, when the quotas started, the Housing Development Board (HDB) did not want to evict households from constrained areas. Hence, to this day, I still observe households above the quotas.

2.4.2 Quota effect on average and group prices

I use a static framework where prices adjust to equate demand and supply. By changing supply differentially across ethnic groups, the quota will likely have opposite

[13]Quotas bind in high quality neighborhoods because these neighborhoods attract many buyers.

effects on buyers of different ethnicities. For example, when the Chinese quotas bind, the Chinese buyers' price increases because they face a decrease in supply (Chinese buyers cannot buy from non-Chinese sellers). I argue that in equilibrium, prices of units in Chinese-constrained areas adjust downwards for non-Chinese buyers so that they will buy into the units that would have been sold to the Chinese buyers, absent the quotas.[14,15]

To see the quota effect on average prices, I compare average prices right above and right below the quotas, assuming that these units are identical except for the quota status. Let the superscript denote buyer type and the subscript denote unit type; $\bar{G}$ and $\underline{G}$ are units above and below the group G quota where G represents (C)hinese, (M)alays and (I)ndians; $v_{\bar{C}}^{NC}$ is the value a non-Chinese buyer has for a unit in a Chinese-constrained area. Buyers of group G have value v with cdf F^G. If group G's ingroup preferences are stronger than the other groups' preferences to live with group G neighbors, we expect that for units in group G-constrained areas, $F^G(v_{\bar{G}})$ will first order stochastically dominate $F^{NG}(v_{\bar{G}})$.We can write the effect of a group G quota on average prices as

$$\Delta \bar{P}_G = \bar{P}_{\bar{G}} - \bar{P}_{\underline{G}} = \frac{P_G^G Q_G^G + P_G^{NG} Q_G^{NG}}{Q_{\bar{G}}^T} - \frac{P_{\underline{G}} \left(Q_{\underline{G}}^G + Q_{\underline{G}}^{NG} \right)}{Q_{\underline{G}}^T} \tag{2.2}$$

where $Q_{\bar{G}}^G$ and $Q_{\bar{G}}^{NG}$ are the number of group G-constrained units bought by group G and non-group-G buyers respectively. Likewise for unconstrained units, $Q_{\underline{G}}^G$ and $Q_{\underline{G}}^{NG}$; Q^T is the total number of units bought by both groups. For simplicity, I assume $Q_{\bar{G}}^T = Q_{\underline{G}}^T$. When the quota binds, prices adjust enough such that the total number of units sold in constrained and unconstrained areas are the same. This implies that only the numerator matters so that revenue changes when average prices change.

[14] Since this is a static framework, I assume the effects of strategic dynamic incentives are second order. An example of such an incentive involves sellers who are right below a quota. If they expect the prices of their units to jump discretely up when the quota binds, they may wait for the quota to bind to sell their unit. These effects could undo any discontinuity in prices caused by the direct quota effect. If I had a panel and I observed frequency of sales, I could test for the presence of these effects. However, even without a panel, I do not find evidence consistent with strategic gaming because the discontinuities in prices at the quotas are fairly large.

[15] Ethnic quotas are related to ethnic prices in two ways. Consider the Chinese quotas. First, Chinese buyers are willing to pay more for Chinese-constrained units because Chinese quota status is positively correlated with Chinese neighborhood quality (Chinese quotas are more likely to bind when the Chinese neighborhood quality is high). Second, Chinese prices increase when the Chinese quotas bind because supply of units to the Chinese buyers is lower. We can think of the first effect as an upward *shift* of the Chinese demand curve relative to the non-Chinese demand curve while the second effect as an upward movement *along* the Chinese buyer's demand curve.

With some algebra, one can show that[16]

$$\Delta \bar{P}_G > 0 \ \text{ if } \frac{1 - F^G\left(P^G_{\bar{G}}\right)}{f^G\left(P_{\underline{G}}\right)} > \frac{1 - F^{NG}\left(P^{NG}_{\bar{G}}\right)}{f^{NG}\left(P_{\underline{G}}\right)} \tag{2.3}$$

The ratios in equation (2.3) represent the trade-offs from changing prices. For example, when the Chinese quota binds (G=C), Chinese buyers have to pay a higher price $\left(P^C_{\bar{C}}\right)$ and non-Chinese buyers pay a lower price $\left(P^{NC}_{\bar{C}}\right)$. This *increases* revenue $\left(\Delta \bar{P}_C > 0\right)$ because $\left[1 - F^C\left(P^C_{\bar{C}}\right)\right]$ households pay a higher price but *decreases* revenue because the higher price implies that $f^C\left(P_{\underline{C}}\right)$ households who would have bought the units absent the quotas would not buy the units now. The second ratio in equation (2.3) has a similar interpretation. When the Chinese quota binds, revenue is higher because more non-Chinese buyers who would not demand these units absent the quota now buy these units $\left(f^{NC}\left(P_{\underline{C}}\right)\right)$, but revenue is lower because$1 - F^{NG}\left(P^{NG}_{\bar{G}}\right)$ non-Chinese households pay a lower price. The net effect on average prices depends on the relative magnitudes of the two ratios.

2.4.3 Regression discontinuity analysis using average prices

I estimate the following equations

$$\begin{aligned}
\ln \bar{P}_{bjit} &= \alpha + \gamma QC_{bjit} + \sum_{k=1}^{4} \phi_k pctC^k_{bji} + \varepsilon_{bjit} && (2.4)\\
\ln \bar{P}_{bjit} &= \alpha + \gamma QC_{bjit} + \sum_{k=1}^{4} \phi_k pctC^k_{bji} + B_{bji}\beta + \tau_t + \omega_i + \varepsilon_{bjit} && (2.5)
\end{aligned}$$

where $\ln \bar{P}_{bjit}$ is the log of the average price of units in block b, neighborhood j, town i and month t; QC_{bjit} is a dummy for whether the Chinese (C) quotas are binding, $pctC^k_{bji}$ are kth order polynomials of the percent of Chinese; B represents other observable characteristics of the block (age of building, number of 1-room units, 2-room units etc.); τ_t and ω_i are month and town fixed effects. I estimate this equation for units that are 10% above and below the Chinese quotas. Similarly, I repeat the

[16]Where $P^G_{\bar{G}}Q^G_{\bar{G}} + P^{NG}_{\bar{G}}Q^{NG}_{\bar{G}} - P_{\underline{G}}\left(Q^G_{\underline{G}} + Q^{NG}_{\underline{G}}\right) = \left(\Delta P^G_{\bar{G}}Q^G_{\bar{G}} + P_{\underline{G}}\Delta Q^G_{\bar{G}}\right) + \left(\Delta P^{NG}_{\bar{G}}Q^{NG}_{\bar{G}} + P_{\underline{G}}\Delta Q^{NG}_{\bar{G}}\right) = \Delta P^G_{\bar{G}}Q^G_{\bar{G}} + \Delta P^{NG}_{\bar{G}}Q^{NG}_{\bar{G}} \cong \frac{1 - F^G\left(P^G_{\bar{G}}\right)}{f^G\left(P_{\underline{G}}\right)} - \frac{1 - F^{NG}\left(P^{NG}_{\bar{G}}\right)}{f^{NG}\left(P_{\underline{G}}\right)}$. $\Delta P^G_{\bar{G}}$ denotes $P^G_{\bar{G}} - P_{\underline{G}}$ (similarly for $\Delta P^{NG}_{\bar{G}}$, $\Delta Q^G_{\bar{G}}$ and $\Delta Q^{NG}_{\bar{G}}$). The first equality is obtained by adding and subtracting $P_{\underline{G}}Q^G_{\bar{G}}$ and $P_{\underline{G}}Q^{NG}_{\bar{G}}$, the second is obtained because $\Delta Q^T_{\bar{G}} = 0$ and the final relationship is obtained because $\Delta P \propto \frac{1}{f(p)}$: The change in demand due to a change in price of Δp is just $\Delta Q = \int_0^{\Delta p} f(p + \varepsilon)\, d\varepsilon = f(p)\,\Delta p$.

analysis for the Malay and Indian quotas. The coefficient of interest is γ, which summarizes the effect of the quota on average prices, at the upper bound. I report results from estimating the discontinuity in average prices at the block quotas.

I assume that the policy rule is perfectly enforced. Since these are public housing units, all resale transactions need to be approved by the HDB. Part of the approval process involves checking that buyers and sellers of a transaction do not violate the ethnic quota rule. It is possible that households sort around the discontinuity. In this case, if households have incentives to undo the discontinuity, then any discontinuity in prices that I estimate would be a lower bound of the actual discontinuity caused by the quotas.

I report results from the regression analysis in Table 2.4. Columns 1-5 correspond to the regression close to the Chinese quota, columns 6-10 correspond to the Malay quota regression and columns 11-15 correspond to the Indian quota regression. For each ethnic quota, I estimate the regression controlling for polynomials of the ethnic proportion, up to the 4th order (first 4 columns) and controlling for observed building characteristics (such as age, number of 1-room flats, number of 2-room flats etc.) and month and town fixed effects (5th column).

Without controlling for building characteristics and fixed effects, average prices are 10% to 11% higher when the Chinese quota binds, but 4% and 3% lower when the Malay and Indian quotas bind. All estimates are significant at the 1% level. In Figure 2.2, I plot the predicted log prices using the estimates from the regression with polynomials of the ethnic proportions up to the 4th order (from columns 4, 9 and 14 in Table 2.4).

These findings are robust to including 1st to 4th order polynomials of the ethnic proportions and for samples restricted to 5% above and below the quotas. Once I add other controls, the estimates for the Indian quotas are not significant. This could be because Indians are such a small minority that almost 95% of the neighborhoods fall within 10% of the Indian quota. Hence, restricting the analysis close to the Indian quotas would essentially be an OLS analysis, since almost all units fall within the 10% window.

2.4.4 Average prices and arbitrage

These findings on average prices are at least consistent with price variation across ethnic groups without perfect arbitrage. One possible concern is inter-temporal arbitrage. For example, in the context of Chinese-constrained units, recall that non-

Chinese sellers cannot sell to Chinese buyers. Sellers could arbitrage away price differences between high-WTP Chinese buyers and low-WTP non-Chinese buyers by waiting for Chinese quotas to become unconstrained, thereby allowing them to sell to the buyer with the highest WTP, instead of (presumably low-WTP) non-Chinese buyers only. Finding that prices are lower when the Malay quotas bind could suggest the absence of perfect arbitrage. Suppose sellers could arbitrage, then, we should expect sellers to be at least as well of as when the quotas were not binding. However, I find that average price (revenue) is lower when the Malay quota binds.

One reason sellers do not arbitrage is waiting costs. If waiting costs were high, then, it would be costly to wait for units to become unconstrained, then sell to the high-WTP buyers. Another reason is that arbitrage is risky if there is lack of coordination. Suppose the Malay quotas were binding at 22% Malays, and, the non-Malay sellers were waiting for the unit to become unconstrained. Then, once the unit becomes unconstrained, there could be a sudden excess supply of units, which may exert a downward force on prices.

2.4.5 Regression discontinuity analysis with estimated group prices

To analyze the estimated group prices (discussed in Section 3.1), notice that knowing the prices that buyers from each ethnic group paid implies knowing the ethnicity of the buyer. In the following specification, I assigned dummy variables for buyer ethnicity and stacked the estimated group prices in a similar regression discontinuity set up

$$\begin{aligned}\ln \hat{P}_{bjit} &= \alpha + \rho_1 QC_{bjit} + \rho_2 QC_{bjit} * QI_{bjit} + \eta_1 buyM_{bjit} + \eta_2 buyI_{bjit} \\ &\quad + \gamma_1 QC_{bjit} * buyM_{bjit} + \gamma_2 QC_{bjit} * buyI_{bjit} + \gamma_3 QC_{bjit} * QI_{bjit} * buyM_{bjit} \\ &\quad + \sum_{k=1}^{4} \phi_k pctC_{bji}^k + B_{bji}\beta + \tau_t + \omega_i + \varepsilon_{bjit} \qquad (2.6)\end{aligned}$$

where now, QC_{bjit} is a dummy that is 1 when only the Chinese quota in block b in neighborhood j, town i and month t is binding; $QC_{bjit} * QI_{bjit}$ is a dummy when both the Chinese and Indian quotas are binding, $buyM_{bjit}$ is a dummy variable that is 1 when the buyer is Malay, $buyI_{bjit}$ is a dummy for an Indian buyer, $pctC_{bji}^k$ is the percent of Chinese in the block, B is a matrix of observable block characteristics, τ_t

and ω_i are the month and town fixed effects, respectively.[17] This equation is estimated for units that are 10% above and below the Chinese block quota. I also estimate a similar equation for the Malay and Indian block quotas. Each time interacting the Malay (Indian) quota, with a dummy for non-Malay (non-Indian) buyers.

The key coefficients of interest here are the $\gamma's$ and the $\rho's$. The idea is to test if Chinese buyers paid a higher price for Chinese-constrained blocks ($\rho s > 0$) and non-Chinese buyers paid a lower price for Chinese-constrained blocks ($\gamma' s \leq 0$). This would indicate that group prices differed for the same neighborhood. Table 2.5 shows the results from the estimation. The 3 columns correspond to the regression for the Chinese, Malay and Indian quotas.

Column 1 shows that Chinese buyers paid 6% more in Chinese-constrained blocks, Chinese and Indian buyers paid 20% more in blocks where both the Chinese and Indian quotas were binding (these estimates of $\rho_1$and ρ_2 are significant at the 1% level). Non-Chinese buyers did not seem to pay a higher price for Chinese-constrained blocks. Moreover, Malay buyers paid 6% less for Chinese- and Indian-constrained blocks (this estimate of γ_3 is significant at the 5% level). So, the signs of the coefficients for the estimated Chinese prices are as expected. The Malay quota has a similar effect on estimated group prices except the coefficients are less significant and I find that blocks where both the Malay and Indian quotas bind, Malay and Indian buyers paid a significantly lower price. The results for the estimation close to the Indian quota (column 3) do not follow the same pattern. Indian buyers paid 4% less (the estimate is significant at the 1% level) when I expected Indian buyers to pay more for Indian-constrained neighborhoods. These findings could again be attributed to the earlier finding that the discontinuity generated by the Indian quotas disappears after controlling for unit characteristics and fixed effects.[18,19]

At this point, one could take the regression discontinuity analysis one step further and use the hedonic method to estimate ethnic ingroup preferences (Rosen, 1974; Chay and Greenstone, 2004). The idea is to address omitted variable bias in demand estimation by examining estimates of the hedonic price function right above and below the discontinuity. This approach comes with two limitations. First, using regression discontinuity design to estimate preferences implies that the estimated

[17]There is no dummy for when both the Malay and Chinese quotas bind because it is impossible to have a block with 87% Chinese and 25% Malays.

[18]In the estimation of the BLP model, since I essentially estimate a separate BLP model for each group (ie. the Indian prices only enter the BLP model for the Indians), this should not affect estimates for the BLP models of the Chinese and the Malays.

[19]The standard errors need to be corrected to account for the fact that the dependent variable is estimated.

taste parameters are only valid for households who are at the discontinuity. There are reasons such as sorting that may suggest that these households are different. A second limitation is that choosing residential neighborhoods is essentially a discrete demand problem while the hedonic method applies to cases with continuous demand. In ongoing work, I explore the possibilities of estimating ethnic preferences using the hedonic model versus the discrete choice model (see Bayer et al., 2007 for a comparison of estimates from hedonic methods and structural methods).

2.5 Utility Specification

To recover ingroup preferences of households away from the discontinuity, I make some assumptions on the functional form of the utility and distributional assumptions on the heterogeneity of individuals. The goal of the structural estimation is to recover individual preferences from aggregate data using the method of simulated moments. I begin with a random coefficients model of individual utility for residential neighborhoods that is then aggregated to obtain market-level demands to be matched with aggregate sample moments.

Suppose we observe m=1, ..., M markets, each with i^G=1, ..., I_m^G buyers of ethnic group G and j=1, ..., J_m neighborhoods. The indirect utility of buyer i of group G from choosing neighborhood j in market m is

$$U_{ijm}^G = X_{jm}^G \beta_i^G - \alpha_i^G P_{jm}^G + \xi_{jm}^G + \varepsilon_{ijm}^G \tag{2.7}$$

for j=1, ..., J_m $\forall m$, where X_{jm}^G is a K-dimensional (row) vector of observed neighborhood characteristics (including ethnic proportions), P_{jm}^G is the price that a buyer of group G has to pay for a unit in neighborhood j in market m, ξ_{jm}^G is the group-specific preference for the unobserved neighborhood attribute and ε_{ijm}^G represents mean-zero, idiosyncratic individual preferences for a buyer of group G, assumed to be independent of neighborhood characteristics, prices and of each other. Note that prices are now indexed by G. This is a consequence of the quotas that I exploit to recover group-specific neighborhood quality (ξ_{jm}^G). To keep notation simple, I will drop the market subscript from here on.

We can write the buyers' taste parameters as a mean component and an individual-specific deviation from the mean

$$\begin{pmatrix} \beta_i^G \\ \alpha_i^G \end{pmatrix} = \begin{pmatrix} \bar{\beta}^G \\ \bar{\alpha}^G \end{pmatrix} + \Sigma \nu_i^G \tag{2.8}$$

where $\nu_{i1}^G, ..., \nu_{iK}^G$ is individual i's unobserved taste for characteristic K, drawn independently (for each individual in each group) from a standard normal distribution and ν_{iP}^G is drawn from a log normal distribution. $\sum$ is a (K+1)x(K+1) dimensional scaling matrix whose diagonal elements are denoted by σ_k and σ_P. Note that I assume mean preferences vary by group but the standard deviation does not ($\sum$ is not indexed by G).

To estimate ethnic preferences, I include the ethnic proportions of the initial stock of owners in neighborhood j. Specifically, the primary neighborhood characteristics of interest are $percent\ Ingroup, percent\ Ingroup^2$ which are the initial percent of ingroups in neighborhood j and its squared. The parameters that represent ethnic preferences are $\bar{\beta}^G_{percent\ Ingroup}$ and $\bar{\beta}^G_{percent\ Ingroup^2}$. For example, $(\bar{\beta}^C_{percent\ Ingroup} + 2\bar{\beta}^C_{percent\ Ingroup^2}\overline{percent\ Chinese} * 0.01$ is the average Chinese buyer's marginal utility to live in a neighborhood with 1% more ingroup members (Chinese) relative to a neighborhood with 1% more outgroup neighbors, evaluated for the households living in the average Chinese neighborhood. Allowing neighborhood quality (ξ) to vary by group allows the interpretation of these parameters as taste for ethnic ingroup interactions, that is separate from the taste for ethnic-specific neighborhood quality. Other observed neighborhood attributes included in X^G_{jm} are school quality, access to public transportation, average number of rooms and average age of HDB blocks in neighborhood j.

The specification is completed with the introduction of an "outside good" (j=0) — buyers may choose not to move.

$$U_{i0}^G = \xi_0^G + \varepsilon_{i0}^G \tag{2.9}$$

Since market shares depend only on differences in utilities, the actual estimation algorithm subtracts U_{i0}^G from U_{ij}^G such that utility is defined relative to the outside good.

Substituting (2.8) into (2.7) and grouping individual-specific terms together, we can write the utility specification parsimoniously as $U_{ij}^G = \delta_j^G + \mu_{ij}^G$ which is simply the mean utility for neighborhood j

$$\delta_j^G = X_j^G \bar{\beta}^G - \bar{\alpha}^G P_j^G + \xi_j^G \tag{2.10}$$

and an individual-specific deviation from that mean

$$\mu_{ij}^G = \sum_k \sigma_k x_{jk}^G \nu_{ik} + \sigma_P P_j^G \nu_{iP}^G + \varepsilon_{ij}^G \tag{2.11}$$

There are two features in (2.10) that will be relevant for estimation. First, utility from neighborhood quality only depends on the quality of that neighborhood alone.[20] This implies that using exogenous attributes of nearby neighborhoods as instruments will satisfy the exclusion restriction. Second, the omitted variable (ξ) enters the mean utility (δ) linearly. This allows the estimation of (2.10) using linear instrumental variable techniques (Berry and Pakes, 2007).

The first two terms in (2.11) are the interaction between consumer tastes and neighborhood characteristics that determine substitution patterns in discrete choice models (McFadden et al., 1977; Hausman and Wise, 1978; BLP, 1995). As heterogeneity (σ) in the unobserved tastes for observed product characteristics increases, neighborhoods that are similar (in characteristics space) become better substitutes.

A potential weakness of the specification is that equation (2.7) has no income. The lack of income data is not particularly helpful for models of residential location choices. In ongoing work, I explore the possibility of proxying for the income of buyers using the income distribution of the entire population.

To return to the model of individual utility, market-level aggregate consumer behavior is obtained by aggregating the choices implied by the individual utility model over the distribution of consumer attributes. Let F_{μ^G} be the population distribution function of individual-level attributes for group G. The fraction of households of group G that choose neighborhood j (aggregate demand) is obtained by integrating over the set of individual attributes that imply a preference for neighborhood j. Let the group G share for neighborhood j be

$$s_j^G(\delta^G, \theta^G; x^G, P^G, F_{\mu^G}) = \int_{A_j^G(\delta^G, \theta^G;\ x^G, P^G)} F_{\mu^G}(d\mu^G), \tag{2.12}$$

where

$$A_j^G(\delta^G, \theta^G;\ x^G, P^G) = \left\{\mu^G : U_{ij}^G > U_{ij'}^G \ \forall\ j' \in J\right\} \tag{2.13}$$

and $\theta^G = \{\bar{\beta}^G, \bar{\alpha}^G, \sigma\}$.

Following the literature, I assume that the idiosyncratic errors, the ε_{ij}^G, have an independently and identically distributed Type I extreme value distribution. This assumption yields the Logit form for the model's choice probabilities. Letting y_i^G denote the choice of buyer i of group G,

[20]This excludes utility specifications, for example, where buyers have higher utility if their neighborhood is better than adjacent neighborhoods.

$$\Pr(y_i^G = j|\nu_i^G, \theta^G, x^G, P^G) = \frac{exp(\delta_j^G + \sum_k \sigma_k x_{jk}^G \nu_{ik} + \sigma_P P_j^G \nu_{iP}^G)}{1 + \sum_{j'} exp(\delta_{j'}^G + \sum_k \sigma_k x_{j'k}^G \nu_{ik} + \sigma_P P_{j'}^G \nu_{iP}^G)} \tag{2.14}$$

Note that the omitted variable (ξ), enters the mean utility term (δ), linearly but it enters demand non-linearly. This complicates the use of standard non-linear instrumental variables method to address omitted variable bias. To address the problem of non-linear omitted variables (ξ), Berry (1994) shows that we can invert the market share function to recover the choice specific constant (δ). Since the choice specific constant is a linear function of quality, once we have recovered δ, we can use standard instrumental variables methods to estimate the mean utility equation (2.10) by finding instruments for endogenous neighborhood attributes that are orthogonal to quality (ξ) .

2.6 Estimation

Using the model of individual utility above, I recover the taste parameters, $\{\bar{\beta}^G, \bar{\alpha}^G, \sigma\}$ by matching aggregate moments predicted from the model to sample moments using the Method of Simulated Moments (MSM).

2.6.1 Method of Simulated Moments

The following moment condition is assumed to hold at the true parameter value, $\theta_0 \in R^p$:

$$E[g(\theta_0)] \equiv E[\xi(\theta_0)| Z] = 0 \tag{2.15}$$

where $g(\bullet) \in R^l$ with $l \geq p$ is a vector of moment functions that specifies that the (structural) error, ξ, is uncorrelated with the instruments, denoted by an JxL matrix, Z. To form the moments, we first need to recover ξ^G.

For each ethnic group G, I first guess values for θ^G which I use to calculate the share function using equations (2.12) and (2.14). I use the contraction mapping provided in Berry (1994) to find the value of δ^G that makes the observed ethnic shares, s^G, equal to the predicted shares defined in equation (2.12). Notice that the integral in the share function, no longer has a closed form solution. I simulate the integral by drawing R=10,000 ν_i^{G}'s independently for each group G and calculating the Logit form in equation (2.14),which is then aggregated out to obtain the market level shares.. After recovering the mean utility, δ^G, by inverting the ethnic share function, I calculate ξ^G using equation (2.10) and the estimated prices. Now that we have an estimate of the (structural) error, ξ^G,we are ready to form the moments.

I stack the moments for the estimation of each ethnic group and define $\theta = \{\theta^C, \theta^M, \theta^I\}$. The simulated moments are

$$\sum_{j=1}^{J} \hat{g}_j(\theta) = \sum_{j=1}^{J} Z_j' \hat{\xi}_j(\theta) \tag{2.16}$$

An MSM estimator, $\widehat{\theta}$, minimizes a weighted quadratic form in $\sum_j^J g_j(\hat{\theta})$:

$$\theta = \arg\min_{\theta \in \Theta} \frac{1}{J} \left[\sum_j^J \hat{g}_j(\theta) \right]' \Omega_J \left[\sum_j^J \hat{g}_j(\theta) \right] \tag{2.17}$$

where Ω_J is an LxL positive, semi-definite weighting matrix. Assume $\Omega_J \xrightarrow{p} \Omega_0$,an LxL positive definite matrix. Define the LxP matrix $G_0 = E[\nabla_{\theta'} g(\theta_0)]$. Under some mild regularity conditions, Pakes and Pollard (1989) and McFadden (1989) show that:

$$\sqrt{J}\left(\hat{\theta} - \theta_0\right) \xrightarrow{d} N(0, \left(1 + R^{-1}\right) * A_0^{-1} B_0 A_0^{-1}) \tag{2.18}$$

where R is the number of simulations, $A_0 \equiv G_0' \Omega_0 G_0$, $B_0 \equiv G_0' \Omega_0 \Lambda_0 \Omega_0 G_0$ and $\Lambda_0 = E\left[g(\theta_0) g(\theta_0)'\right] = Var\left[g(\theta_0)\right]$. If a consistent estimate of Λ_0^{-1} is used as the weighting matrix, the MSM estimator, $\hat{\theta}$, is asymptotically efficient, with its asymptotic variance being $Avar(\hat{\theta}) = (1 + R^{-1}) * \left(G_0' \Lambda_0^{-1} G_0\right)^{-1} / J$.

To account for the error from using estimated prices instead of actual prices, I follow the discussion in Newey (1984) on sequential estimators and method of moments. We can think of the exercise above in the context of GMM where there are 2 moments and parameters are estimated sequentially. First, we estimate prices with parameters π and γ (the first moments are $g^1(\pi, \gamma)$). Then, using these parameters as inputs, the second moments are just the standard BLP moments, $g^2(\theta, \hat{\pi}, \hat{\gamma})$. To calculate the standard errors, I stack the BLP moments and the moments from the estimation of prices and calculate standard errors using equation (2.18).

2.6.2 Instruments

In this section, I discuss the instruments for ethnic proportions and price. Note that the exclusion restriction holds by definition for most of the instruments discussed below because utility from neighborhood j only depends on the characteristics of neighborhood j that are in the utility function.

Ethnic proportions

I instrument for ethnic proportions using attributes of nearby neighborhoods as well as historical settlement data. I use the sum of the exogenous characteristics of rival products (in my case, I use attributes of neighborhoods in 1-3km rings, 3-5km rings and 5-7km rings). The attributes include average number of rooms, average age of building, school quality and average distance to the closest subway station. I chose 1km as the cutoff because the neighborhoods would be far enough to avoid spatial correlation with own neighborhood. I chose 2km widths so that all neighborhoods would have at least one nearby neighborhood within the ring.[21] The idea of using attributes of nearby neighborhoods is that Chinese, Malays and Indians have different preferences for neighborhood attributes, perhaps due to demographics such as family sizes. Forty-three percent of Malay households have 5 or more family members while only 24% and 26% of Chinese and Indian households have such large families (HDB, 2000). The thought experiment involves 2 similar neighborhoods where one is surrounded by neighborhoods with many big units and the other is surrounded by neighborhoods with few big units. Malay households would tend to sort into the neighborhood surrounded by neighborhoods with few big units since many of these large Malay households will prefer big units.

In addition to using attributes of nearby neighborhoods to instrument for ethnic proportions, I also use a dummy variable on whether units are to the east of the early Malay settlement. The idea is that exogenous assignment of Malay settlements to the east of the Singapore River imply that subsequent Malay neighborhoods were more likely to develop on the east side of the Singapore River.

Group-specific prices

I instrument for group-specific prices using exogenous characteristics of nearby neighborhoods and estimated quotas. Specifically, I follow BLP (1995) and use the sum of the exogenous characteristics of rival neighborhoods. Attributes of nearby neighborhoods are valid instruments for prices if markets are competitive.

While attributes of nearby neighborhoods could instrument for common prices, it is hard to think of them as good instruments for group-specific prices. One would expect ethnic quotas to be highly correlated with group-specific prices (Chinese prices

[21]One neighborhood, Changi Village, is located at the Eastern tip of Singapore. There are no neighborhoods within 1-3km of Changi Village. I assign values of the instruments to be zero for Changi Village.

are high when the Chinese quota binds). However, actual quotas are not valid instruments because they are positively correlated with the structural error term, ξ_j^G in equation (2.10). Therefore, I use estimated quotas where the quotas are estimated using only exogenous variables. This ensures that variation from the estimated quotas only derives from exogenous variation in neighborhood characteristics.

To estimate quotas, I exploit the step function in the policy rule. First, I estimate the block (neighborhood) level ethnic proportions using own block (neighborhood) characteristics and nearby neighborhood characteristics. Then, I assign the estimated block (neighborhood) quotas to be 1 if the estimated block (neighborhood) ethnic proportions are above the block (neighborhood) level quotas and 0 otherwise. The estimation equation for block and neighborhood proportions are

$$\begin{aligned} pctG_{bj} &= \gamma_0 + X_{bj}^{ex}\gamma_1 + \gamma_2 East_j + Z_j\gamma_3 + u_{bj} && (2.19) \\ pctG_j &= \rho_0 + X_j^{ex}\rho_1 + \rho_2 East_j + Z_j\rho_3 + v_j && (2.20) \end{aligned}$$

where G=(C)hinese, (M)alays and (I)ndians, b indexes blocks and j indexes neighborhoods. The variable, $pctG$ is the percent of residents from group G, X^{ex} is the set of exogenous observed characteristics, $East$ is a dummy for whether the neighborhood is to the east of the early Malay settlements, Z is the set of exogenous characteristics of nearby neighborhoods.

Constructing instruments using this step function is akin to the regression discontinuity identification strategy. Quotas affect prices according to the step function (above the upper bound, prices differ across groups). The instruments, X^{ex}, Z and $East$, could affect prices but the effect is not discontinuous at the upper bound. Therefore, even though the quotas were estimated by projecting neighborhood proportions onto the space of X^{ex}, Z and $East$, the estimated quotas should still have power to predict group prices. To test this, I estimate the following equations

$$QG_j = \chi\widehat{QG}_j + v_j \qquad (2.21)$$

$$QG_j = \phi_0 + X_j^{ex}\phi_1 + \phi_2 East_j + Z_j\phi_3 + \phi_4\widehat{QG}_j + \omega_j \qquad (2.22)$$

That is, I regress the actual quota status for Chinese, Malay and Indian quotas (QG_j) on the estimated quota status $\left(\widehat{QG}_j\right)$. For example, QC_j is the percent of blocks in neighborhood j where the Chinese quota is binding in a month.[22] Also, I regress actual

[22]This is a percentage instead of a dummy because there are block and neighborhood quotas. This number is 1 when the neighborhood quota binds in a month (all blocks are constrained) and less than 1 when some blocks are hitting the block quota. For the dependent variable, I use data from

quotas on the estimated quotas,controlling for the full set of exogenous variables. If quotas have power above and beyond the exogenous characteristics used to estimate them, then, the coefficient, ϕ_4, should be significant. This regression is akin to the first stage of an instrumental variables regression except the dependent variable is not group-specific prices (what the quotas instrument for) because I do not observe group-specific prices.

2.7 Results

Table 2.6 reports results from regressions of actual quotas on estimated quotas. The first 3 columns do not control for other exogenous instruments that I used to predict ethnic proportions. The coefficients on the estimated quotas are all positive and significant. After controlling for exogenous instruments (columns 4-6), Chinese and Malay estimated quotas remain significant but Indian quotas are negative and not significant. This could be because the discontinuity from the Indian quotas essentially disappears after controlling for fixed effects and unit characteristics as shown in the regression discontinuity results (the last column in Table 2.4). The idea of using estimated quotas to instrument for prices relies on the step function of the quotas. Since the step function from the Indian quotas disappears after adding controls, it is not surprising that the coefficient on the estimated Indian quota is insignificant after adding controls.

Table 2.7 reports results from estimating a Logit model with OLS (columns 1-3) and IV (columns 4-6) where the dependent variables are the log of the ethnic shares, $\ln(s_{jm}^G)$ subtracted by the log of the ethnic share for the outside good, $\ln\left(s_{0m}^G\right)$ and G indexes for the (C)hinese, the (M)alays and the (I)ndians. I use estimated group-specific prices in this regression. Most of the coefficients are of the right sign in OLS but this does not mean estimates are not biased due to omitted variables.

Table 2.8 reports results from estimating the random coefficients model using group-specific prices for 170 neighborhoods. The top panel reports results on the mean of the taste parameters, $\bar{\beta}$ and $\bar{\alpha}$ and the bottom panel reports results on the heterogeneity term, σ. The first column refers to estimates that are restricted to be common across groups and the next three columns are preference parameters for Chinese, Malays and Indians.

Interpreting the magnitudes of the taste parameters, living in a neighborhood where the average building is 10 years older is as bad as living 2.3 km further away

March 2005, the earliest month in my dataset that is relevant for this group of movers.

from the subway station. These parameters are significant at the 1% and 5% level respectively. On average, households prefer the outside good (not moving) since the marginal utility of the constant term is significantly negative. There is substantial heterogeneity in the taste for rooms. The coefficient on price enters negatively but it is not significant. Most coefficients enter with the right sign except the marginal utility for rooms which is negative and significant.

All groups want to live with at least some members of their own group. The Chinese and Indians have ingroup preferences that are inverted U-shaped such that in neighborhoods where there are enough members of their ingroup, Chinese and Indians prefer neighbors from other groups, on the margin (the parameters are significantly positive for *percent Ingroup* but significantly negative for *percent* $Ingroup^2$). The estimated marginal utilities for Malays are positive for both terms, albeit not significant.

Using the estimates on marginal utilities from Table 2.8, I calculate the marginal rates of substitution (MRS) between *percent Ingroup* and distance to the subway station as well as the MRS between *percent Ingroup* and the average age of a building (where older is considered worse). Because of the quadratic term on *percent Ingroup*, the MRS changes when the ethnic proportions in a neighborhood change. The MRS is calculated as the marginal utility for a neighborhood with a 1 standard deviation increase in the percent of ingroup members divided by the marginal utility for a neighborhood that is 1km closer to the closest subway station.[23] I calculate the MRS with respect to age of buildings in a similar way.

Figure 2.3 plots the MRS with respect to the distance to the subway station as a function of the neighborhood ethnic proportion. Each point on the line is an MRS. The plots show that Chinese and Indian ingroup preferences are inverted U-shaped because the marginal utilities for ingroup neighbors are positive below 43% Chinese and 8% Indians respectively but are negative for neighborhoods with a higher concentration of ingroup members.

Table 2.9 shows the MRS's evaluated at neighborhoods with the mean, the 10th percentile, and the 90th percentile of ethnic proportions. The MRS for the Indians (the 2nd and 3rd panels in the last column) suggests that Indians have ingroup preferences that are inverted U-shaped. The average Indian household living in a

[23]Since distance to subway and age of building are bad characteristics, I used the negative of their marginal utilities in the MRS calculation so that a positive MRS reflects ingroup preferences and a negative MRS reflects outgroup preferences for the marginal neighbor. For example, the MRS with respect to distance to the subway is calculated as $\frac{\left(\bar{\beta}^{C}_{pctIngroup}+2\,\bar{\beta}^{C}_{pctIngroup^2}\,\overline{pctChinese}\right)*0.01}{-\bar{\beta}_{subway}}$

neighborhood with 5% Indians (the 10th percentile) is willing to substitute to a neighborhood that is 2.56km *further* from the closest subway station in exchange for a neighborhood with a 1 standard deviation (3%) increase in the proportion of Indians. This distance is relatively far considering the average household reported that the maximum acceptable walking distance to a subway station is 0.53km (HDB, 2000).[24] On the contrary, the average Indian household living in a neighborhood with 8% Indians (the mean) is willing to substitute to a neighborhood that is 0.21km *closer* to the subway station for the same 1 standard deviation increase in the percent of ingroup neighbors.

Although the Chinese have ingroup preferences that are inverted U-shaped as shown in Figure 2.3, the MRS's in Table 2.9 for the Chinese (the first column) are all negative. This is because the estimates suggests that the Chinese have strong ingroup preferences for at least 43% of Chinese but once a neighborhood has more than 43% Chinese, the average Chinese household desires a neighbor from the outgroups. Since all neighborhoods in my sample have more than 43% Chinese (the minimum is 61%), the MRS for the Chinese evaluated at all neighborhoods in the sample are negative. The Malays who live in a neighborhood with an average percent of Malays (13%), are willing to substitute to a neighborhood that is 2.1km further away from subway stations to live in a neighborhood with a 1 standard deviation (7%) increase in the percent of Malay neighbors, although the estimates for the Malays are not significant.

The finding that Chinese exhibit positive outgroup preferences at the margin is suggestive evidence against ethnic discrimination. The concern with ethnic discrimination is that Malay enclaves may form even when Malays have no ingroup preferences, simply because the Chinese are discriminating against Malays and forcing them into Malay enclaves. This could mean that my estimate of Malay ingroup preferences is an overestimate. However, to the extent that the Chinese are discriminating against outgroups, the Chinese should have strong negative outgroup preferences, which is not what I find.

[24]This is based on a survey conducted by the HDB to study the profile of public housing residents in Singapore. The distance is calculated by the surveyors in HDB based on a rate of conversion of 10 minutes to 500m of walking distance. There are 2 types of subway stations; the Mass Rapid Transit (MRT) is the primary subway while the Light Rapid Transit (LRT) is relatively new and limited. The number given above refers to the distance to an MRT station.

2.8 Conclusion and Future Research

In this paper, I build and estimate a discrete choice model of residential location choices by combining policy variation akin to a regression discontinuity framework with a structural demand estimation framework, motivated by the ethnic housing quotas in Singapore. I exploit the step function in the quota rule as well as within neighborhood, across group variation in prices to identify ethnic preferences from correlated tastes for ethnic neighborhood quality.

I find that all groups have strong preferences to live with at least some other members of their ethnic group. However, the Chinese and the Indians exhibit preferences that are inverted U-shaped so that after a neighborhood reaches 43% Chinese and 8% Indians respectively, they would rather add a new neighbor from the other group.

In ongoing work, I explore the possibilities of combining hedonic methods and regression discontinuity to estimate ethnic preferences as well as to compare estimates from the hedonic method and the discrete choice method. Future work will also include the use of estimates on ethnic preferences in this paper to simulate counterfactuals and estimate deadweight losses from integration policies that will need to be weighed against social benefits of integration. The challenge in performing such welfare calculations is that sorting models typically have multiple equilibria. Findings from this type of simulation can inform the relative deadweight losses and the distributional implications of various integration policies.

Bibliography

Ananat-Oltmans, Elizabeth. 2007. "The Wrong Side(s) of the Tracks: Estimating the Causal Effects of Racial Segregation on City Outcomes." NBER Working Paper No. 13343.

Angrist, Josh and Victor Lavy. 1999. "Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement." *Quarterly Journal of Economics*, 114(2): 533-75.

Athey, Susan and Guido Imbens. 2007. "Discrete Choice Models with Multiple Unobserved Choice Characteristics." Forthcoming, *International Economic Review.*

Bajari, Patrick and Matthew E. Kahn. 2005. "Estimating Housing Demand With an Application to Explaining Racial Segregation in Cities." *Journal of Business and Economic Statistics*, 23(1): 20-33.

Bayer, Patrick and Christopher Timmins. 2007. "Estimating Equilibrium Models of Sorting Across Locations." *The Economic Journal*, 117(518): 353-374.

Bayer, Patrick, Fernando Ferreira and Robert McMillan. 2007. "A Unified Framework for Measuring Preferences for Schools and Neighborhoods." NBER Working Paper No. 13236.

Bayer, Patrick, Robert McMillan, and Kim Rueben. 2004. "An Equilibrium Model of Sorting in an Urban Housing Market." NBER Working Paper No. 10865.

Berry, Steven T. 1994. "Estimating Discrete-Choice Models of Product Differentiation." *The Rand Journal of Economics*, 25(2): 242–62.

Berry, Steven T., James Levinsohn, and Ariel Pakes. 1995. "Automobile Prices in Market Equilibrium." *Econometrica,* 63(4): 841–90.

Berry, Steven T., James Levinsohn, and Ariel Pakes. 2004. "Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market." *Journal of Political Economy,* 112(1): 68–105.

Brennan, Brian. 2000. "Background on MTO." *Moving To Opportunity Research.* http://www.wws.princeton.edu/~kling/mto/background.htm

Card, David, Alexandre Mas and Jesse Rothstein. 2007. "Tipping and the Dynamics of Segregation." NBER Working Paper No. 13052.

Chay, Kenneth and Michael Greenstone. 2004. "Does Air Quality Matter? Evidence from the Housing Market." *Journal of Political Economy*, 113(2): 376-424.

Crawfurd, John. 1828. "Journal of an Embassy from the Governor-General of India to the Courts of Siam and Cochin-China : Exhibiting a View of the Actual State of those Kingdoms." URL: http://dlxs.library.cornell.edu/s/sea/index.php (Cornell University, Southeast Asia Visions)

Cutler, David M.; Glaeser, Edward L. 1997. "Are Ghettoes Good or Bad?" *Quarterly Journal of Economics*, 112(3): 827-72.

Cutler, David M.; Glaeser, Edward L.; Vigdor, Jacob L. 1999. "The Rise and Decline of the American Ghetto." *Journal of Political Economy*, 107(3): 455-506.

Das, S., G. Steven Olley, and Ariel Pakes. 1995. "The Market for TVs." Yale Working Paper.

Hausman, Jerry A., and David A. Wise. 1978. "A Conditional Probit Model for Qualitative Choice: Discrete Decisions Recognizing Interdependence and Heterogeneous Preferences." Econometrica 46 (March): 403–26.

Housing Development Board (HDB). 2000. *Profile of Residents Living in HDB Flats.* Singapore: HDB.

Housing Development Board (HDB). 2006. *Annual Report.* Singapore: HDB.

Lum, Sau Kim and Tan Mui Tien. 2003. "Integrating Differentials: A Study on the HDB Ethnic Integration Policy." *Crestimes*, Singapore Center for Real Estate (CRES), 3(1): 5.

Manski, Charles F. 1993. "Identification of Endogenous Social Effects: The Reflection Problem." *Review of Economic Studies,* 60 (3): 531-42.

McFadden, Daniel. 1974. "Conditional Logit Analysis of Qualitative Choice Behavior." In *Frontiers in Econometrics*, edited by Paul Zarembka. New York: Academic Press.

McFadden, Daniel, Antti Talvitie and Associates. 1977. *Demand Model Estimation and Validation.* Berkeley: The Institute of Transportation Studies, University of California, Berkeley and Irvine.

McFadden, Daniel. 1989. "A Method of Simulated Moments for Estimation of Discrete Response Models without Numerical Integration," *Econometrica*, 57(5): 995-1026.

Newey, Whitney. 1984. "A Method of Moments Interpretation of Sequential Estimators." *Economic Letters*, 14: 201-206.

Pakes, Ariel, and David Pollard. 1989. "Simulation and the Asymptotics of Optimization Estimators." *Econometrica,* 57(5): 1027–57.

Schelling, Thomas C. 1971. "Dynamic Models of Segregation." *Journal of Mathematical Sociology,* 1(3): 143-186. (Abbreviated version appeared as "Models of Segregation." in the American Economic Review, 59(2), May 1969).

Singapore Department of Statistics. 2006. "Singapore Resident Population, 1990 - 2006."

Singapore Department of Statistics, Singapore 2000 Census. "Key Indicators of the Resident Population."

Virtual Map Online Street Directory. http://www.streetdirectory.com/

Table 2.1: Neighborhood and Block Level Quotas Relative to National Ethnic Proportions

	Neighborhood Quotas	Block Quotas	National Proportion (2000)
Chinese	84%	87%	77%
Malay	22%	25%	14%
Indian	10%	13%	8%

Source: 2000 Census (Singstat), Lum and Tan (2003)

Table 2. 2: The Relationship between Quotas, Buyer Ethnicity, Seller Ethnicity, and Prices

Binding Quota	Buyer Ethnicity	Seller Ethnicity	Status	Group-Specific Prices
Chinese	Chinese	Chinese	Allowed	High Chinese Prices
	Non-Chinese	Non-Chinese	Allowed	
	Non-Chinese	Chinese	Allowed	
	Chinese	Non-Chinese	Not Allowed	
Malay	Malay	Malay	Allowed	High Malay Prices
	Non-Malay	Non-Malay	Allowed	
	Non-Malay	Malay	Allowed	
	Malay	Non-Malay	Not Allowed	
Indian	Indian	Indian	Allowed	High Indian Prices
	Non-Indian	Non-Indian	Allowed	
	Non-Indian	Indian	Allowed	
	Indian	Non-Indian	Not Allowed	

Notes: The lnk between group-specific prices and quotas is premised on two correlations: (i) Chinese prices are high when the Chinese neighborhood quality is high; (ii) Chinese quotas are more likely to bind in neighborhoods with high Chinese quality because these neighborhoods attract relatively more Chinese. These two correlations imply that Chinese prices are likely to be positively correlated with Chinese quotas. In addition, if we assume the Chinese, Malay and Indian quality are not perfectly correlated, then, the prices that Chinese buyers are willing-to-pay would be higher than Malay and Indian prices in Chinese-constrained neighborhoods, and lower than Malay and Indian prices in Malay- and Indian-constrained neighborhoods.

Table 2.3: Summary Statistics

Variable	N	Mean	Std. Dev.	Description
Chinese Share	170	0.09%	0.11%	Percent of Chinese in a market who chose a neighborhood
Malay Share	170	0.13%	0.14%	Percent of Malays in a market who chose a neighborhood
Indian Share	170	0.30%	0.31%	Percent of Indians in a market who chose a neighborhood
Price	170	239,888	50,769	Average transaction price in a neighborhood (Singapore dollars)
Chinese Neighborhood Quota	170	0.08	0.25	Percent of months Chinese neighborhood quota binds
Malay Neighborhood Quota	170	0.05	0.19	Percent of months Malay neighborhood quota binds
Indian Neighborhood Quota	170	0.17	0.33	Percent of months Indian neighborhood quota binds
Chinese Block Quota	170	0.10	0.18	Percent of months and blocks Chinese block quota binds
Malay Block Quota	170	0.05	0.12	Percent of months and blocks Malay block quota binds
Indian Block Quota	170	0.09	0.15	Percent of months and blocks Indian block quota binds
Chinese Quota	170	0.18	0.29	Percent of months and blocks any Chinese quota binds
Malay Quota	170	0.11	0.23	Percent of months and blocks any Malay quota binds
Indian Quota	170	0.25	0.35	Percent of months and blocks any Indian quota binds
Percent Chinese	170	79%	7%	Percent of Chinese in a neighborhood
Percent Malay	170	13%	7%	Percent of Malays in a neighborhood
Percent Indian	170	8%	3%	Percent of Indians in a neighborhood
School Quality	170	3.15	4.21	Total number of awards received by schools in a neighborhood
Subway	170	0.80	0.55	Distance to the closest subway station
Rooms	170	4.12	0.63	Number of rooms in an average flat in the neighborhood
Age	170	19.22	7.11	Average age of HDB blocks in the neighborhood

Note: School quality is measured as the total number of awards given to primary, secondary schools and tertiary institutions by the Singapore Ministry of Education.

Table 2.4: Regression Discontinuity Results on Average Prices

	Dependent variables														
	Ln Price	Ln Price	Ln Price	Ln Price	Ln Price	Ln Price	Ln Price	Ln Price	Ln Price	Ln Price	Ln Price	Ln Price	Ln Price	Ln Price	Ln Price
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Chinese Quota	0.10***	0.10***	0.11***	0.11***	0.06***										
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)										
Percent Chinese	-0.17***	1.70	328.45***	219.38***	42.51										
	(0.06)	(2.03)	(50.25)	(33.43)	(50.71)										
Percent Chinese^2		-1.10	-383.21***	-192.06***	-36.90										
		(1.20)	(58.72)	(29.30)	(43.82)										
Percent Chinese^3			148.63***	dropped	dropped										
			(22.84)												
Percent Chinese^4				43.27***	8.21										
				(6.63)	(9.68)										
Malay Quota						-0.04***	-0.04***	-0.04***	-0.04***	-0.03***					
						(0.01)	(0.01)	(0.01)	(0.01)	(0.01)					
Percent Malay						0.10	-1.73***	-8.15**	-77.81***	-3.10					
						(0.06)	(0.51)	(3.49)	(21.28)	(22.82)					
Percent Malay^2							4.01***	32.13**	488.59***	11.61					
							(1.10)	(15.16)	(138.35)	(147.89)					
Percent Malay^3								-39.77*	1,338.48**	-15.08					
								(21.39)	(391.85)	(416.34)					
Percent Malay^4									1,354.56**	1.15					
									(408.09)	(429.84)					
Indian Quota											-0.03***	-0.03***	-0.03***	-0.03***	-0.01
											(0.01)	(0.01)	(0.01)	(0.01)	(0.02)
Percent Indian											-0.28***	-0.11	2.32***	-10.64***	1.71
											(0.06)	(0.23)	(0.72)	(2.02)	(1.81)
Percent Indian^2												-0.85	-25.42***	177.01***	-16.35
												(1.09)	(6.94)	(30.25)	(25.18)
Percent Indian^3													73.11***	1,189.35**	45.84
													(20.38)	(184.76)	(145.45)
Percent Indian^4														2,691.83**	-7.05
														(391.55)	(303.64)
Controls	N	N	N	N	Y	N	N	N	N	Y	N	N	N	N	Y
Month	N	N	N	N	Y	N	N	N	N	Y	N	N	N	N	Y
Town	N	N	N	N	Y	N	N	N	N	Y	N	N	N	N	Y
Obs	14136	14136	14136	14136	14136	11471	11471	11471	11471	11471	23871	23871	23871	23871	23871
R-squared	0.01	0.01	0.02	0.02	0.74	0.004	0.01	0.01	0.01	0.71	0.003	0.004	0.004	0.01	0.72

Standard errors in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Note: The first 5 columns are restricted to be 10% above and below the Chinese quota, the next 5 columns, correspond to the Malay quota, and the last 5 columns correspond to the Indian quota. Columns 1-4, 6-9, 11-14 include results from regressions with a quota dummy and polynomials (up to 4th order) of the block level ethnic proportions. Columns 5, 10 and 15 control for unit characteristics (average age of building, its squared, number of 1-room flats, 2-room flats etc.), month and town fixed effects.

Table 2.5: Regression Discontinuity Results on Estimated Group-Specific Prices

	Dependent variables		
	Predicted Ln Price	Predicted Ln Price	Predicted Ln Price
	(1)	(2)	(3)
Chinese Quota	0.06***		
	(0.01)		
Chinese Quota * Indian Quota	0.20***		
	(0.01)		
Chinese Quota * Malay Buyer	-0.02**		
	(0.01)		
Chinese Quota * Indian Buyer	-0.003		
	(0.01)		
Chinese Quota * Indian Quota * Malay Buyer	-0.06***		
	(0.02)		
Malay Quota		0.03***	
		(0.01)	
Malay Quota * Indian Quota		-0.07***	
		(0.01)	
Malay Quota * Indian Buyer		0.01	
		(0.01)	
Malay Quota * Chinese Buyer		0.01	
		(0.01)	
Malay Quota * Indian Quota * Chinese Buyer		-0.01	
		(0.01)	
Indian Quota			-0.04***
			(0.005)
Indian Quota * Chinese Quota			0.18***
			(0.01)
Indian Quota * Malay Quota			-0.09***
			(0.01)
Indian Quota * Chinese Buyer			-0.005
			(0.01)
Indian Quota * Malay Buyer			0.002
			(0.01)
Indian Quota * Chinese Quota * Malay Buyer			-0.03**
			(0.01)
Indian Quota * Malay Quota * Chinese Buyer			-0.02
			(0.01)
Chinese Buyer		0.004	0.01***
		(0.004)	(0.003)
Malay Buyer	-0.01***		-0.004
	(0.004)		(0.003)
Indian Buyer	-0.01	-0.01	
	(0.004)	(0.003)	
Controls	Y	Y	Y
Obs	10767	10149	17394
R-squared	0.24	0.35	0.28

Standard errors in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Note: Each column is a regression restricted to 10% above and below the Chinese, Malay and Indian quotas. Controls include average age of building, its squared, number of 1-room flats, 2-room flats etc. The omitted group is the Chinese buyer (column 1), the Malay buyer (column 2) and the Indian buyer (column 3).

Table 2.6: Regression of Actual Quota Status on Estimated Quota Status

	Dependent variables					
	Actual Chinese Quota	Actual Malay Quota	Actual Indian Quota	Actual Chinese Quota	Actual Malay Quota	Actual Indian Quota
	(1)	(2)	(3)	(4)	(5)	(6)
Predicted Chinese Quota	0.48***			0.17*		
	(0.05)			(0.07)		
Predicted Malay Quota		0.65***			0.42***	
		(0.12)			(0.12)	
Predicted Indian Quota			0.30**			-0.14
			(0.10)			(0.13)
Controls	N	N	N	Y	Y	Y
Obs	170	170	170	170	170	170
Fstat	85.63	31.50	8.61	7.23	3.82	2.33
R-squared	0.34	0.16	0.05	0.46	0.31	0.22

Standard errors in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Note: Results in columns 1-3 do not include other controls that will be in the full model, columns 4-6 control for own attributes and the instruments. The instruments are the sum of school awards, the distance to the closest subway station, the average age of buildings, the average number of rooms for nearby neighborhoods within 1-3km, 3-5km and within 5-7km, as well as a dummy for being in the east of the early Malay settlements.

Table 2.7: Logit and IV Logit

	Dependent variables					
	Ln Chinese Share	Ln Malay Share	Ln Indian Share	Ln Chinese Share	Ln Malay Share	Ln Indian Share
	(1)	(2)	(3)	(4)	(5)	(6)
School Quality	0.12***	0.09***	0.11***	0.12***	0.09***	0.10***
	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)	(0.03)
Distance to Subway	-0.42**	-0.14	-0.39**	-0.42**	-0.09	-0.36*
	(0.19)	(0.16)	(0.18)	(0.21)	(0.20)	(0.22)
Average No. of Rooms	-0.08***	-0.05***	-0.08***	-0.09***	-0.05**	-0.04
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)
Average Age of Buildings	-0.09	-0.26	-0.24	-0.77	-0.85*	-0.01
	(0.30)	(0.24)	(0.27)	(0.48)	(0.46)	(0.40)
Percent Ingroup	76.90***	17.54***	22.68**	97.32*	26.87**	28.34
	(24.81)	(5.05)	(9.25)	(52.71)	(12.25)	(22.75)
Percent Ingroup^2	-51.40***	-34.78**	-86.67**	-67.61**	-59.27	-230.05**
	(15.77)	(17.14)	(40.13)	(34.09)	(39.15)	(107.16)
Price	-11.72***	-4.68**	-6.63***	-0.76	5.71	-1.14
	(2.85)	(2.14)	(2.42)	(6.64)	(7.06)	(4.75)
Constant	-27.58***	-2.29*	-0.89	-33.35*	-3.13**	-3.43*
	(9.73)	(1.16)	(1.33)	(20.15)	(1.50)	(1.94)
Obs	170	170	170	170	170	170
R-squared	0.37	0.38	0.27	0.29	0.29	-0.06

Standard errors in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Notes: The dependent variable is the log of the Chinese share for neighborhood j subtracted by the log of the Chinese share for the outside good (columns 1and 4). I define the dependent variables for the other columns in a similar manner. The instruments are the sum of school awards, the distance to the closest subway station, the average age of buildings, the average number of rooms for nearby neighborhoods within 1-3km, 3-5km and within 5-7km, a dummy for being in the east of the early Malay settlements as well as the estimated quotas.

Table 2.8: Random Coefficients Logit

Variables	Units	Common Taste Parameters (1)	Chinese Taste Parameters (2)	Malay Taste Parameters (3)	Indian Taste Parameters (4)
Means ($\bar{\beta}, \bar{\alpha}$)					
Constant		-5.69***			
		(1.70)			
School Quality	.1 awards	1.46***			
		(0.15)			
Distance to Subway	1 km	-0.24**			
		(0.13)			
Average No. of Rooms	0.1 rooms	-6.77***			
		(2.34)			
Average Age of Buildings	0.01 years	-5.63***			
		(1.43)			
Percent Ingroup			4.83**	4.00	5.64**
			(2.70)	(7.85)	(1.46)
Percent Ingroup^2			-5.58**	1.17	-3.62**
			(2.96)	(2.71)	(1.09)
Price	S$million	-2.11			
		(3.29)			
Heterogeneity (σ)					
Constant		-1.30***			
		(0.23)			
Average No. of Rooms	0.1 rooms	-3.78***			
		(0.77)			
Price	S$million	0.29			
		(0.39)			

Standard errors in parentheses adjusted for sequential estimators, using Newey (1984)
* significant at 10%; ** significant at 5%; *** significant at 1%

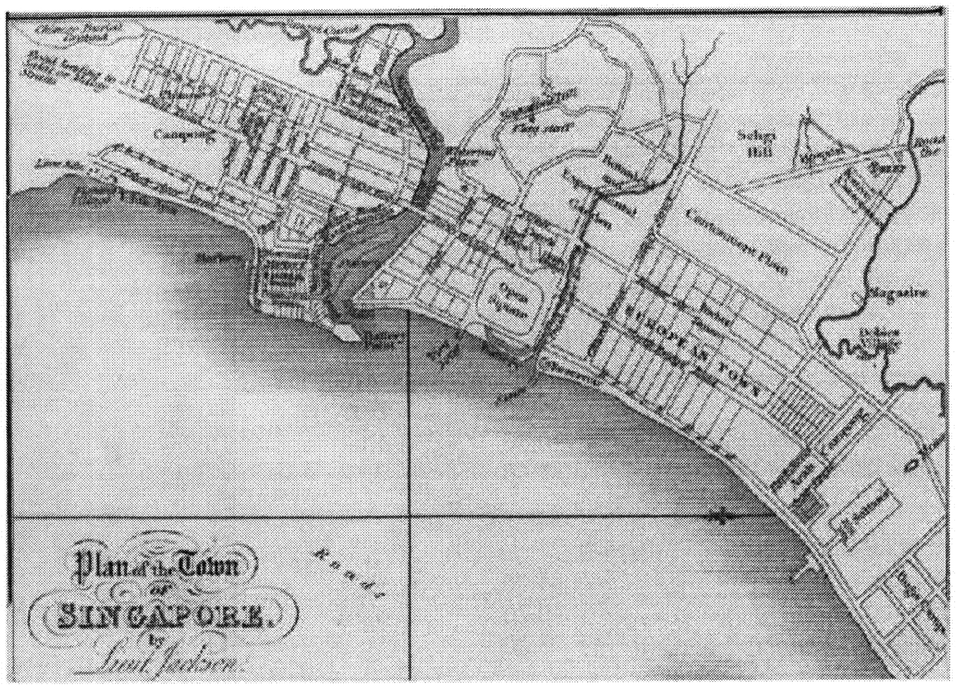
Notes: Variables are scaled so that the mean is between 0 and 1. The units are in the table. For example, the coefficient on *School Quality* implies that an increase in 10 awards is associated with an increase of 1.46 utils. For the variable *Percent Ingroup*, *percent Chinese*, *percent Chinese*^{2} and *percent Malay* are not scaled; *percent Malay*^{2} and *percent Indian* are multiplied by 10; *percent Indian*^{2} is multiplied by 100. The instruments are the sum of school awards, the distance to the closest subway station, the average age of buildings, the average number of rooms for nearby neighborhoods within 1-3km, 3-5km and within 5-7km, a dummy for being in the east of the early Malay settlements as well as the estimated quotas.

Table 2.9: MRS Evaluated at Various Ethnic Proportions in the Sample

	Chinese (1)	Malays (2)	Indians (3)
Relevant statistics for ethnic proportions			
Mean of *Percent Ingroup*	79%	13%	8%
10th percentile of *Percent Ingroup*	70%	5%	5%
90th percentile of *Percent Ingroup*	88%	22%	12%
Standard Deviation of *Percent Ingroup*	7%	7%	3%
MRS relative to distance to subway (km)			
MRS at mean of *Percent Ingroup*	-1.18	2.09	-0.21
MRS at 10th percentile of *Percent Ingroup*	-0.88	1.54	2.56
MRS at 90th percentile of *Percent Ingroup*	-1.48	2.72	-3.90
MRS relative to age of building (years)			
MRS at mean of *Percent Ingroup*	-4.95	8.76	-0.86
MRS at 10th percentile of *Percent Ingroup*	-3.70	6.43	10.73
MRS at 90th percentile of *Percent Ingroup*	-6.20	11.38	-16.32

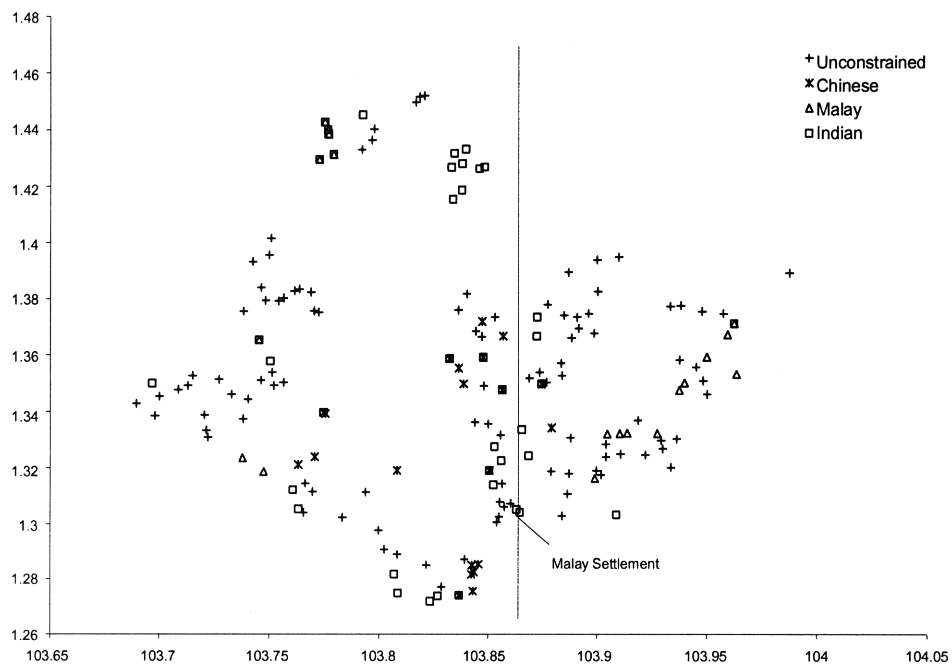
Note: This table shows calculations of the MRS evaluated at different ethnic proportions. The top panel shows the relevant statistics for the ethnic proportions. The second panel represents the MRS relative to the distance to the subway station and the third panel represents the MRS relative to the average age of the building. Since ingroup preferences are quadratic in percent ingroup, the marginal utilities vary with the percent of ingroups in the neighborhood. Each number in the 2nd and 3rd panel represents the marginal rate of substitution evaluated at neighborhoods with the mean, the 10th percentile, and the 90th percentile of *Percent Ingroup*. The MRS is calculated as the estimated marginal utility to live in a neighborhood with a one standard deviation increase in percent ingroup divided by the (negative of the) marginal utility for distance to the closest subway station (2nd panel) as well as the (negative of the) marginal utility for the average building age (3rd panel). Since distance to the subway and building age are both bad attributes, I use the negative of their marginal utilities in the denominators so that a positive MRS reflects ingroup preferences and a negative MRS reflects outgroup preferences. Although the MRS's for the Chinese are negative for all values of *percent Chinese* calculated for this table, Chinese do have positive MRS's when the percent of Chinese in a neighborhood is less than 43% (see Figure 2.3). However, all neighborhoods in the sample have more than 43% Chinese, which is why the MRS shown here for the Chinese are all negative.

Figure 2.1a: Map of EthnicSettlements in Early 19th Century Singapore



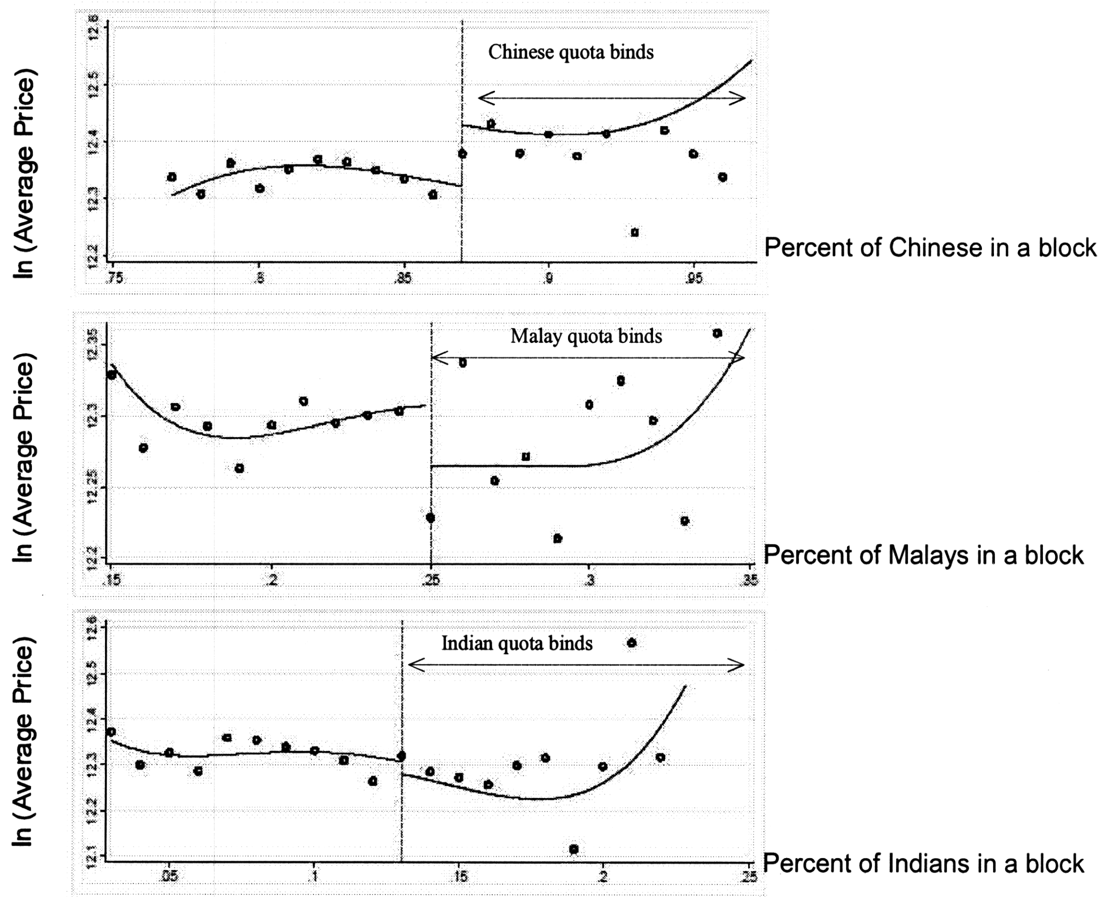
Source: Crawford (1828)
Notes: The Malay settlements ("Arab Kampong" and "Bugis Kampong") are in the south east corner of the map, just east of the European Town. The Chinese and Indian areas are to the west of the Singapore River.

Figure 2.1b: Map of Current Neighborhood by Quota Status



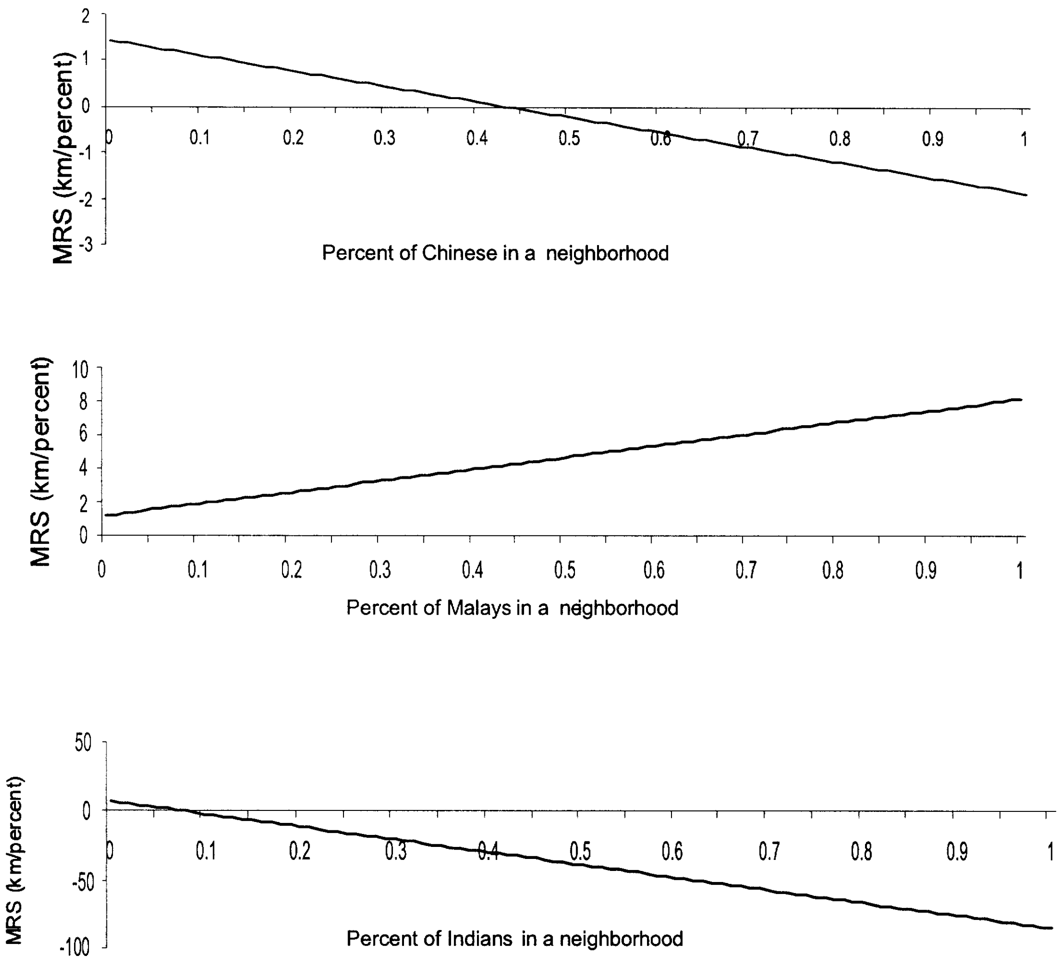
Source: Virtual Map Online Street Directory
Note: This is a map of 170 neighborhoods comprising unconstrained neighborhoods, Chinese-, Malay- and Indian - constrained neighborhoods. The line indicates the eastern tip of the early Malay settlements.

Figure 2.2: Average Transaction Prices above and below Chinese, Malay and Indian Quotas



Notes: Each panel in this figure is constructed using the following procedure for observations within 10% of the ethnic quotas: (i) regress the log of average transaction prices on the corresponding running variable (to a 4th order polynomial) and a dummy that is one when the corresponding block quota is binding; (ii) plot the predicted prices above and below the quota separately (iii) plot means of ln(price) for each 1% bin. I repeat the exercise for the Malay quotas and Indian quotas.

Figure 2.3: MRS between *Percent Ingroup* and Distance to the Subway Station for the Chinese, Malays and Indians as a F unction of Neighborhood Ethnic Proportions



Note: Each point on the line represents the ratio of the estimated marginal utility to live in a neighborhood with a one standard deviation increase in *Percent Ingroup* divided by the marginal utility for distance to the closest subway station. Since ingroup preferences are quadratic in *Percent Ingroup*, the marginal utility varies with the percent of ingroup in the neighborhood (x-axis). The plot shows that in neighborhoods with less than 43% Chinese and 8% Indians respectively, the Chinese and Indians have a preference for a marginal neighbor who is Chinese. Above 43%, Chinese have a preference for a marginal neighbor who is from the outgroup. The minimums and maximums of the ethnic proportions in the sample are 61% and 98% for the Chinese; 0.6% and 33% for the Malays; 1.3% and 26% for the Indians.

Data Appendix

In this section, I describe some variables in more detail and list the corresponding data sources.

Choice data

I match the postal codes of individuals in the 2005 and the 2006 phonebook. Movers have to update their contact information within a month of moving. Households can request for phone and address records to be unlisted at a charge of $20 per annum plus a one-time administrative fee of $20. The phone company updates the data every year on April 1st. For my dataset, I assume movers moved between April 2005 and March 2006 and they changed their phone records immediately after they move.

Neighborhoods

I use six-digit postal codes to define neighborhoods. Blocks that are within the same sector (defined by the first 2 digits of the postal code) and whose 3 digit block numbers share the same first digit are assigned to the same neighborhood.

School quality

I obtain data on awards given to primary, secondary schools and tertiary institutions from the Singapore Ministry of Education website. The school quality is defined as the total number of awards received from all schools and tertiary institutions in a neighborhood.

Access to subway

For each neighborhood, I calculate the distance (in kilometers) from the midpoint of the neighborhood to the closest Mass Rapid Transit (MRT) or Light Rapid Transit (LRT) station using latitude and longitude data obtained from a popular local online street directory, http://www.streetdirectory.com/.

Age

This is obtained from the resale transactions data on the HDB website. Since all blocks in the resale market were sold at some point in my dataset, I observe the age of each HDB block. I use the average age of HDB blocks in a neighborhood.

Rooms

I purchased this data from the HDB. For each HDB block, I have the number of type 1 flats, type 2 flats etc. There are 8 types of HDB flats comprising 1-room to 5-room flats, executive flats, HUDC and multi-generational flats. 1-room flats are studios, 2-room flats are 1 bedroom flats and so on. Executive flats, HUDC and multi-generational flats are defined as 6-room flats in my dataset.

Quotas

I collected monthly data on the ethnic quotas from the public HDB website, beginning March 2005. These are dummy variables for whether a block was constrained. If all blocks were constrained in a neighborhood, I say the neighborhood quota is binding.

Chapter 3

The Welfare Consequences of Ethnic Housing Quotas in Singapore

3.1 Introduction

Quotas are a major policy tool to encourage integration across groups of individuals, whether the groups be gender, ethnicities or nationalities. These quotas vary from affirmative action quotas in Malaysian universities, to hiring quotas in the US police force, quotas on immigrants and the ethnic housing quotas in Singapore. Economists know very little about the welfare consequences of these quotas partly because we lack information on how much individuals prefer to be with own types. In this paper, I use findings from Chapter 2 to answer 2 questions: What does the first best look like in the Singapore housing market? Do the ethnic housing quotas in Singapore get us closer to the first best?

Externalities provide economic rationale for public policies such as quotas, but extermalities also complicate welfare analysis. An individual's decision on where to locate affects ethnic distributions, in aggregate. Assuming households care about ethnicities of their neighbors, individual location choices can affect the utility of neighbors. Housing prices do not internalize these externalities and externalities suggest that the decentralized equilibrium may not be optimal. Therefore, standard welfare analysis using prices and market valuation only is not sufficient. We need to know the shapes of preferences for neighbors' ethnicities to determine what is first best.

Can quotas achieve the first best? Schelling (1971) conjectured that ethnic housing quotas could be a useful policy tool in cases with multiple equilibria where there are both segregated and also mixed equilibria. Suppose the market started in a segregated equilibrium, such as Singapore did in 1989, and a mixed equilibrium exists, then the idea is that ethnic housing quotas could actually move the equilibrium from a segregated one to the desired mix equilibrium. This paper investigates this conjecture.

What does the first best look like? I find that the first best has fewer Chinese- and Indian-segregated neighborhoods but more Malay-segregated neighborhoods. Comparing data from 3 segregated towns before the quota, I find that after 10 years since the introduction of the quota, the decentralized equilibrium had moved the Malay and Indian proportions significantly closer (within 5%) to first best.

3.2 Theoretical framework[1]

The social planner's problem is to find the allocation of ethnicities into neighborhoods that will maximize a utilitarian social welfare function. Individual i's utility from living in neighborhood j is:

$$U_{ij}^G = \delta_j^G + \mu_{ij}^G \tag{3.1}$$
$$\delta_j^G = X_j^G \overline{\beta}^G + \xi_j^G \tag{3.2}$$
$$\mu_{ij}^G = \sum_k \sigma_k x_{jk}^G \nu_{ik}^G + \varepsilon_{ij}^G \tag{3.3}$$

where G represents (C)hinese, (M)alay and (I)ndian, X_j^G is a 1xk matrix of observed neighborhood characteristics, ξ_j^G is the unobserved ethnic-specific neighborhood quality (ethnic amenities, for example), υ_{ik}^G and ε_{ij}^G are the idiosyncratic taste shocks for characteristic k and neighborhood j respectively, $\overline{\beta}$ and σ are the parameters of the taste distribution (mean and standard deviation). U_{ij}^G is a combination of group G's mean utility for neighborhood j, δ_j^G, and the individual-specific deviation from the mean, μ_{ij}^G. The observed characteristics include a *constant, percent of ingroup,* $(percent\ of\ ingroup)^2$, *school quality, number of rooms, distance to the closest subway station,* and *age of buildings.*

The social planner is assigning individuals into neighborhoods. Let $j(i)$ represent the assignment function where $j(i) = j'$ means that individual i is assigned to neighborhood j'. Using this notation, individual i's utility is defined as $U_i \equiv U_{ij(i)}$ and $U_i = U_{ij'}$ if $j(i) = j'$.[2] The social planner's problem is to choose the assignment of individuals that maximizes a utilitarian social welfare function.

[1] The details of the policy are described in Chapters 1 and 2.

[2] The definition $U_i \equiv U_{ij(i)}$departs from the decentralized equilibrium where the individual chooses the neighborhood that maximizes his utility, $U_i = \max_j U_{ij}$.

$$\max_{\{j(i)\}_{i=1}^{i=I}} \sum_{G=C,M,I} \sum_{i} U_i^G = \max_{\{j(i)\}_{i=1}^{i=I}} \sum_{G=C,M,I} \sum_{i} U_{ij(i)}^G \, (percent\ ingroup) \tag{3.4}$$

The key is that assignment of ethnicities into neighborhoods changes the aggregate ethnic distribution in neighborhoods, *percent ingroup*, which affects everyone's utility, not just the individual making the decision. In a decentralized equilibrium, individuals choose the neighborhood that maximizes his own utility, without internalizing the effect of his choice on *percent ingroup*, which affects the utility of his neighbors. The social planner, by contrast, may assign an individual to a neighborhood that is suboptimal for him, but optimal for his neighbors.

There are two key factors that determine what the first best ethnic distribution looks like. First, the relative sizes of the ethnic groups. The Chinese are the majority (77%) and Malays and Indians are the minority (14% and 9%). Because the Malays are such a small minority, it is not possible to have all neighborhoods be 100% Malays even if this was preferred by Malays.

Secondly, the shapes of the preferences of each ethnic group will also determine what the first best looks like. Table 3.1 reports estimated taste parameters from Chapter 2.[3] The relevant marginal utilities for the average Chinese, Malay and Indian are:

$$\frac{\partial \delta^C}{\partial\, (percent\ Chinese)} = 4.83 - 11.16 * (percent\ Chinese) \tag{3.5}$$

$$\frac{\partial \delta^M}{\partial\, (percent\ Malay)} = 4 + 23.4 * (percent\ Malay) \tag{3.6}$$

$$\frac{\partial \delta^I}{\partial\, (percent\ Indian)} = 56.4 - 724 * (percent\ Indian) \tag{3.7}$$

On average, all ethnic groups want neighbors from their own group (the first term in the marginal utility), but the average Chinese and the average Indian do not want too many of their own group. Once a neighborhood reaches 43% Chinese and 8% Indian, respectively, the average Chinese and Indian prefers a neighbor from other ethnicities. The average Indian has the strongest preference for diversity because he

[3] In Chapter 2, note that $(percent\ Malay)^2$, *percent Indian*, $(percent\ Indian)^2$ are multiplied by 10, 10 and 100 respectively so that all variables in the preference estimation simulations were scaled to be between 0 and 1. I account for this when calculating the marginal utilities using the estimated taste parameters.

has the steepest indifference curve (for a 1% increase in the percent of Indian, the average Indian utility changes by the most, compared to a 1% increase in the percent of Chinese and the percent of Malay).

In a decentralized equilibrium, holding all other characteristics fixed, the average Malay living in a neighborhood with more than 43% Chinese will choose to move out of a Chinese neighborhood into a Malay neighborhood because he prefers to live with his own type. However, this imposes a negative externality on his Chinese neighbors, on average, because at this point, an increase in *percent Chinese*, introduces a negative marginal utility to the average Chinese. The social planner could internalize this externality by assigning some Malays to live with Chinese.

Relative to the ethnic distribution now, I expect the first best density of *percent Chinese* and *percent Indian* will have more mass on the left (more neighborhoods with low Chinese and Indian proportions). In particular, the density of *percent Indian* will have the most mass on the left since the average Indian has the strongest preference for diversity. By contrast, the first best density of *percent Malay* will have more mass on the right relative to the ethnic distribution now because they have strong preferences for own group neighbors.

Although the optimal mix for the average Chinese and Indian household are 43% and 8% respectively, not all neighborhoods will converge towards these percentages for 2 reasons. First, it is impossible to have all neighborhoods with only 43% Chinese because Chinese make up 77% of the population. Secondly, there is heterogeneity in other characteristics, in particular, the ethnic-specific amenities $\left(\xi_j^G\right)$.

3.3 Simulation

I use the same data for the 170 neighborhoods, as described in Chapter 2. The idea of the simulation is to find the ethnic distributions, $\{percent\ Chinese_j, percent\ Malay_j\}_{j=1}^{J}$ that maximize a utilitarian social welfare function, using the following steps:

1. For each market m, find the number of neighborhoods, J_m in that market.
2. Assume each neighborhood has 100 units. Let the population in that market be $N_m = 100 * J_m$. The number of Chinese, Malays and Indians are $0.77N_m$, $0.14N_m$ and $0.09N_m$ respectively.

3. Randomly draw J_m Chinese and Malay proportions, $\{percentChinese_j, percent\ Malay_j\}_{j=1}^{J_m}$ such that the mean Chinese proportion in the market is 77% and the Chinese and Malay proportions sum to less than 1 for each neighborhood.

4. Assign Chinese, Malays and Indians to live in each neighborhood, where the number assigned to each neighborhood is determined using the Chinese and Malay proportions drawn in the previous step. This step determines the function $j(i)$ for each i.

5. Draw the corresponding idiosyncratic taste shocks for each ethnic group, where the individual taste for characteristics, v_{ik}^{G}, is common across neighborhoods and the logit error, ε_{ij}^{G}, is neighborhood-specific.

6. Calculate utility, $U_i = U_{ij(i)}$ where the assignment of individual i to neighborhood j is determined in step 4.

7. Sum the individual utilities to get the social welfare function.

8. Repeat steps 3-7 10,000 times for each market. Determine the ethnic proportions that maximize a utilitarian social welfare function.

3.4 Results

Figure 3.1 plots the density of *percent Chinese*, *percent Malay* and *percent Indian* now (dashed line) and under first best (solid line). Recall that once they have enough of their own ethnic groups, the Chinese and Indians start prefering a marginal neighbor from other groups, but the Malays do not. As expected, there are more neighborhoods that have low Chinese and Indian proportions and more neighborhoods with higher Malay proportions. Currently, there are no neighborhoods with less than 60% Chinese but in the first best case, there are 25 such neighborhoods. Similarly, there are currently 91 neighborhoods with fewer than 8% Indians but in the first best case, there are 113 such neighborhoods. For the Malays, there are 21 more neighborhoods with more than 33% Malays in the first best case. Figure 3.2 maps the changes in ethnic proportions. Both the increase (white bubbles) and decrease (black bubbles) in ethnic proportions relative to the current proportions is uniformly distributed geographically, except for the Malays. Most of the changes in Malay proportions is concentrated to the East where the neighborhoods tend to have higher Malay-specific amenities (ξ_j^M is high). Table 3.2 shows some percentiles of the current density and the first best case.

While both Chinese and Indians have inverted U-shaped preferences for own ethnic group neighbors, the density for the Chinese has a thin left tail under first best while the Indian case has significantly more mass on the left. This is partly because Indians have a stronger preference for diversity (*preference effect*) and they constitute a small minority of the population (*size effect*) so that it is possible to have many neighborhoods with low Indian proportions. In Figure 3.3, I investigate how much the difference in group sizes matters by running the simulations assuming all ethnic groups are of equal size (I remove the size effect and just have the preference effect). Now, the Chinese and Indian densities have similar shapes. This suggests that the size effect is significant. With the size effect, since the Chinese are the majority, even though the average Chinese living in a neighborhood with more than 43% Chinese prefers a marginal non-Chinese neighbor, there can never be too many mixed neighborhoods. Note that the scale of the y-axes in Figure 3.3 are different. There are 69 neighborhoods with less than 20% Chinese but there are 100 neighborhoods with less than 20% Indians. This difference reflects the preference effect (Indians have stronger tastes for diversity. Thus, there are more neighborhoods with low Indian proportions even after removing the size effect).

While the first best has more diverse neighborhoods, Table 3.3 shows that the first best also has more neighborhoods that are above the quotas. Under first best, 71% more neighborhoods would be Chinese-constrained, almost twice as many neighborhoods would be Malay-constrained and 33% more neighborhoods would be Indian-constrained. Since Malays prefer their own type only, the first best has more neighborhoods with a higher Malay proportion and it is not surprising that more neighborhoods will be Malay-constrained. For the Chinese, to create more diverse neighborhoods means some Chinese have to be moved into other neighborhoods, thus, increasing the Chinese proportion in those neighborhoods. While we have 25 more neighborhoods with less than 60% Chinese, this also means 36 more neighborhoods that would have to be above the Chinese quota.

Do the ethnic housing quotas get the market closer towards first best? Table 3.4 looks at 3 towns where there is data on ethnic proportions before the quota was implemented in 1989. Redhill was known as a Chinese town, Bedok was a Malay town and Yishun was an Indian town. Twenty years ago, in 1988, the Malay and Indian proportions in Bedok and Yishun were almost 4 times the first best levels. Ten years after the introduction of the quotas, in 1998, the Malay and Indian towns, Bedok and Yishun, were already within 5% of the first best Malay and Indian proportions.

The quotas seem to be successful at moving the Malay and Indian proportions closer towards first best but not the Chinese proportions, probably because the Chinese have preferences for diversity but they are such a majority that it is hard to lower Chinese neighborhood proportions.

3.5 Conclusion

Quotas are a major policy tool to encourage integration across groups of individuals. Externalities provide economic rationale for public policies such as quotas, but extermalities also complicate the use of prices to do welfare analysis because prices do not internalize these externalities. This paper uses results on estimated preference parameters from Chapter 2 to simulate the first best equilibrium and compare it to the existing equilibrium with the ethnic housing quotas in Singapore.

I find that the first best has fewer Chinese- and Indian-segregated neighborhoods but more Malay-segregated neighborhoods, consistent with the shapes of preferences estimated in Chapter 2. Comparing data from 3 segregated towns before the quota, I find that after 10 years since the introduction of the quota, the decentralized equilibrium had moved the Malay and Indian proportions significantly closer to first best.

There are two major caveats to the discussion above. First, the estimation in Chapter 2 assumes that all ethnic groups share a common taste for characteristics such as school quality. Second, this welfare exercise assumes that other characteristics, such as school quality, do not change in response to the change in ethnic proportions. If Chinese proportions are positively correlated with school quality, creating more diverse neighborhoods by lowering Chinese proportions could lower the school quality in those neighborhoods. If Chinese care more about education than the minorities, this effect could off-set the Chinese taste for diversity since lowering Chinese proportions would also lower school quality.

Bibliography

Housing Development Board (HDB). 2000. *Profile of Residents Living in HDB Flats.* Singapore: HDB.

Schelling, Thomas C. 1971. "Dynamic Models of Segregation." *Journal of Mathematical Sociology,* 1(3): 143-186. (Abbreviated version appeared as "Models of Segregation." in the American Economic Review, 59(2), May 1969).

Singapore Straits Times, 7 January 1989 (Table 3.4).

Table 3.1: Estimated Taste Parameters

Variables	Units	Common Taste Parameters (1)	Chinese Taste Parameters (2)	Malay Taste Parameters (3)	Indian Taste Parameters (4)
Means $(\bar{\beta}, \bar{\alpha})$					
Constant		-5.69***			
		(1.70)			
School Quality	.1 awards	1.46***			
		(0.15)			
Distance to Subway	1 km	-0.24**			
		(0.13)			
Average No. of Rooms	0.1 rooms	-6.77***			
		(2.34)			
Average Age of Buildings	0.01 years	-5.63***			
		(1.43)			
Percent Ingroup			4.83**	4.00	5.64**
			(2.70)	(7.85)	(1.46)
Percent Ingroup^2			-5.58**	1.17	-3.62**
			(2.96)	(2.71)	(1.09)
Price	S$million	-2.11			
		(3.29)			
Heterogeneity (σ)					
Constant		-1.30***			
		(0.23)			
Average No. of Rooms	0.1 rooms	-3.78***			
		(0.77)			
Price	S$million	0.29			
		(0.39)			

Standard errors in parentheses adjusted for sequential estimators, using Newey (1984)
* significant at 10%; ** significant at 5%; *** significant at 1%

Notes:Obtained from Table 8 of Chapter 2 . Variables are scaled so that the mean is between 0 and 1. The units are in the table. For example, the coefficient on *School Quality* implies that an increase in 10 awards is associated with an increase of 1.46 utils. For the variable *Percent Ingroup*, *percent Chinese*, *percent Chinese*^{2} and *percent Malay* are not scaled; *percent Malay*^{2} and *percent Indian* are multiplied by 10; *percent Indian*^{2} is multiplied by 100.

Table 3.2: Percentiles of Ethnic Distributions, First Best and Now

	10th percentile	25th percentile	Median	75th percentile	90th percentile
Percent Chinese, First Best	36%	73%	85%	93%	97%
Percent Chinese, Now	69%	74%	80%	85%	88%
Percent Malay, First Best	1%	3%	6%	17%	53%
Percent Malay, Now	5%	7%	12%	18%	22%
Percent Indian, First Best	1%	2%	5%	5%	17%
Percent Indian, Now	5%	7%	8%	10%	12%

Table 3.3: Number of Neighborhoods Quota-Constrained, First Best and Now

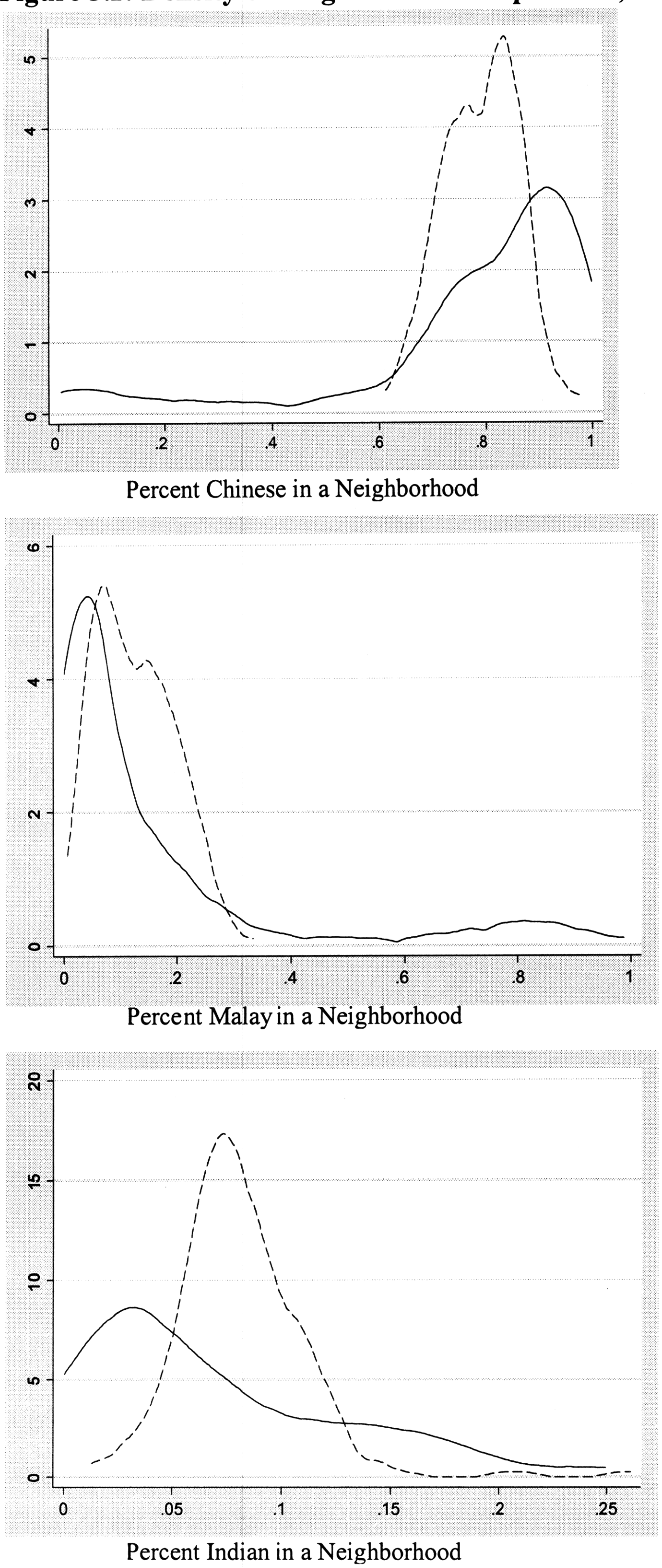
	Chinese	Malay	Indian
Now	51	17	36
First Best	87	33	48

Table 3.4: Ethnic Proportions of Three Towns, Before and After the Quota, First Best

	Before (1988)	After (1998)	First Best
Percent Chinese in Redhill	87%	84%	75%
Percent Malay in Bedok	59%	19%	15%
Percent Indian in Yishun	24%	11%	6%

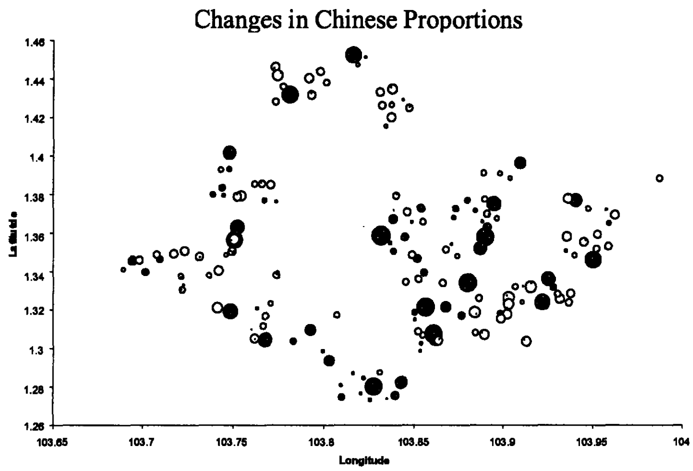
Source: Straits Times 7 January 1989, HDB profile of residents in HDB flats, 1998.

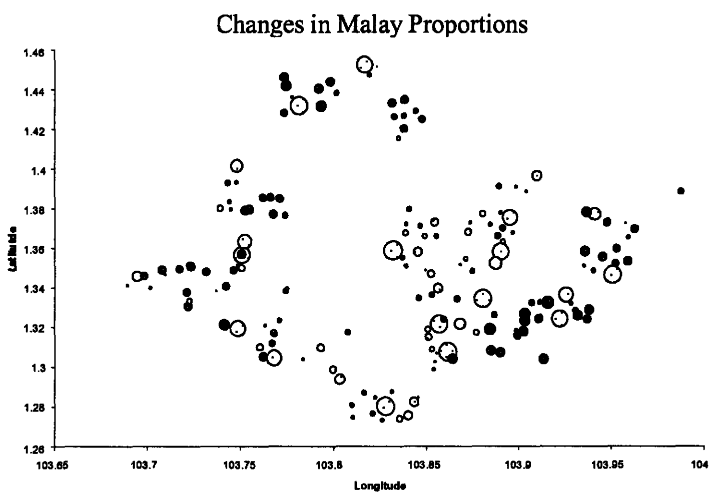
Figure 3.1: Density of Neighborhood Proportions, First Best and Now

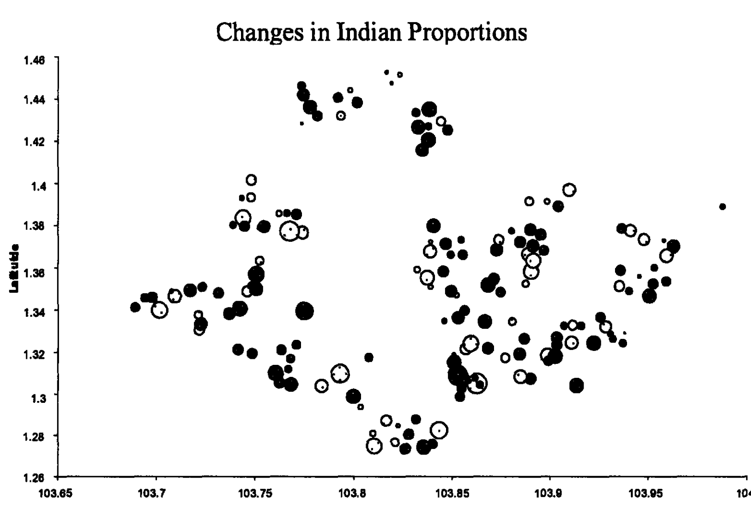


Notes: Dashed lines represent the density now (- - - -); Solid lines (——) represent the first best density.

Figure 3.2: Map of Changes in Neighborhood (First Best – Now)

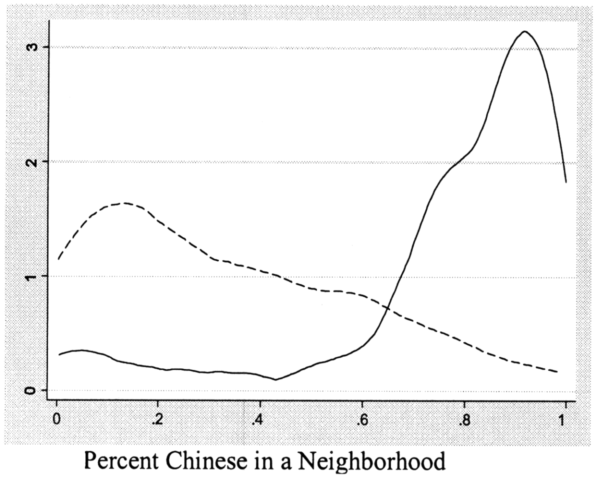


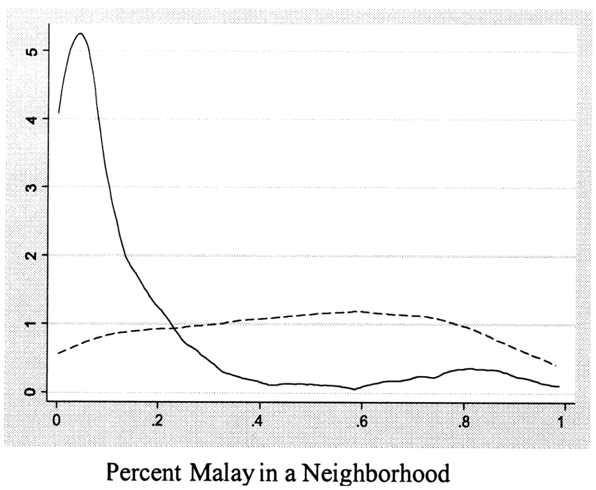


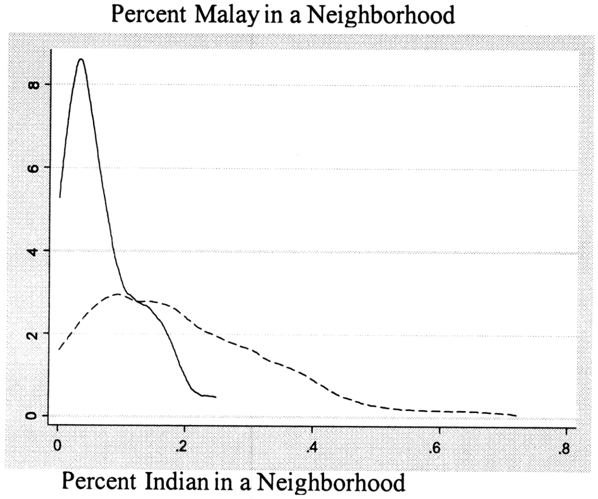


Notes: Each bubble is a neighborhood. A black bubble means the neighborhood proportion in first best is lower compared to now. A white bubble represents an increase in neighborhood proportion in the first best relative to current proportions. The size of the bubble corresponds to the size of the change.

Figure 3.3: Density of Neighborhood Proportions, First Best. Assuming Equal and Actual Population Sizes for Ethnic Groups







Notes: Dashed lines represent the first best density without the size effect, assuming the population sizes are the same for all three ethnic groups (----); Solid lines (——) represent the first best density with the size effect, ie. using actual population sizes (from Figure 3.1).