LANGUAGE, IMMIGRATION, AND CITIES

by

QIANG LI

B.Eng., Tongji University, 2000
M.Econ., Tongji University, 2003

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES

(Business Administration)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

November 2008

Abstract

This dissertation analyzes the complex relationships between language, immigration, and labor and housing market outcomes. First, I model the urban labor market as segmented by language barriers. The prediction of this segmentation theory is confirmed by Canadian Census data, which allow me to identify a worker's labor market segment by her work language. Second, I explore whether the housing market reflects people's willingness to pay for higher quality social-ethnic interactions. By combining housing transaction data and Census information, I am able to test such a relationship with positive results. Finally, I ask what properties housing price series have if some people have better knowledge of the future immigration/migration flows to a city. Under this setup, the price series become serially correlated and the price volatility varies over time. The model also explains the long-standing price-volume relationship in housing transaction data.

Table of Contents

Appendices

List of Tables

List of Figures

Acknowledgements

I am deeply indebted to my thesis supervisor Professor Bob Helsley for his continued support and guidance throughout my graduate program. This dissertation is impossible without the numerous discussions that prompted me to improve the chapters again and again. Professor Tsur Somerville acted as my graduate advisor for the first two years. His help and suggestions are invaluable to me. He also provided many valuable comments to improve my dissertation. Professor Sanghoon Lee helped me a lot on my presentation skills and often acted as a constructive critic for my research since he arrived in UBC in 2005. The comments from Professor David Green for the labor market segmentation chapter greatly improved the structure and content of the chapter. I also want to thank Professor Jim Brander and Professor Keith Head for their many good suggestions.

Dedication

TO MY PARENTS

Chapter 1

Introduction

Urban economists have long recognized the importance of social interactions in the formation and development of cities. Jacobs (1969) discusses how firms and workers interact in cities and the primary function of these interactions in the continuous prosperity of cities. In her other classic, Jacobs (1961) discusses the importance of constant social interaction among individuals for the livelihood of a neighborhood.

Language is probably the most important factor that affects social interactions. On the one hand, the common knowledge of a language facilitates our daily communications. On the other hand, language affects the development of culture identities among individuals. These identities then affect people's preferences toward whom to interact with and how.

I explore the complex relationships between language, immigration, and cities in this dissertation. In chapter 2, I ask how wages are determined in a multicultural city. The labor market is segmented by language barriers across cultures. I test this segmentation theory using Canadian Census data. The special feature in the data allows me to use the work language to identify a worker's labor market. The predictions of the model are confirmed by the data.

In chapter 3, I ask whether neighborhood housing price is related to its composition of language groups. Different neighborhoods provide different potential social interaction quality. Those with higher quality will command higher prices. I derive some conditions under which we should expect such a relationship. I combine the housing transactions and Census data for the metropolitan area of Vancouver. Using this dataset, I confirm the existence of such a relationship.

In chapter 4, I build a theory of speculation in the housing market when some people have better knowledge of future migration flows to a city. For cities that experience population growth through migration, the model provides predictions about the properties of housing prices. Speculation can explain the serial correlation in the returns of housing prices. In addition, speculation also induces volatility clustering, or the variation of price volatility, and positive correlation between price and transaction volume.

Chapter 2

Language and Labor Market Segmentation

2.1 Introduction

> *And the whole earth was of one language, and of one speech. ... And the Lord came down to see the city and the tower, which the children of men builded. And the Lord said, 'Behold, the people is one, and they have all one language; and this they begin to do: and now nothing will be restrained from them, which they have imagined to do.'*
> (Genesis 11:1-6)

Language is an important part of our lives. Our common knowledge about the words and rules of a language facilitates our daily communications. In this chapter, I study whether language affects workers' labor market outcomes and more importantly how it exerts its influence. I present a labor market segmentation theory in which language plays a central role. Figure 2.1 shows an illustration of this theory.

Workers in a city differ in two important respects. They belong to either the majority language group, which has a larger population, or the minority language group. They also choose to enter either the majority labor market or the minority labor market. The language groups are differentiated by individuals' home languages, while the labor market segments are distinguished by the languages commonly used in business communications.[1]

Two types of wage gaps exist. As Figure 2.1 shows, the first type exists between majority workers and minority workers who work in the majority labor market. This gap is defined as Within-Labor-Market Wage Gap. In the majority market, a minority worker can communicate in the majority language, but some tacit language barriers hinder her ability to communicate

[1] Notice in Figure 2.1 'majority workers in the minority market' has been crossed to indicate that those workers may not exist in reality. See Section 2.2.5 for details.

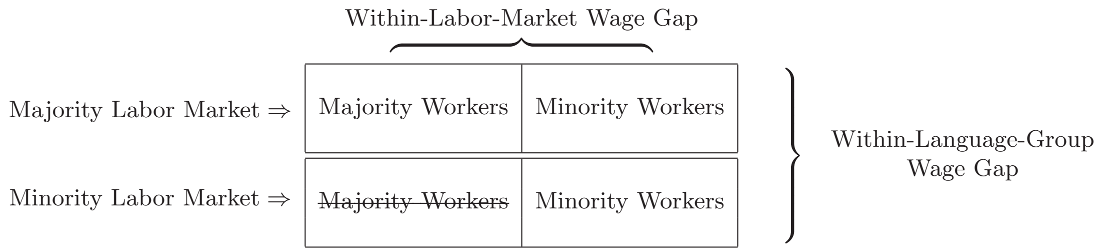

Figure 2.1: Language and Labor Market Segmentation

effectively with her coworkers and manager.[2] As a result, she is not as productive as a majority worker, so she earns less.

The second type of wage gap exists between minority workers who work in different labor market segments. I define this as Within-Language-Group Wage Gap. In the model, a worker's wage depends on the quality of the match between her job and her skill. The majority labor market has a larger number of workers and firms, which means it can offer a higher matching quality and a higher wage for workers. This market thickness effect has been modeled by Helsley and Strange (1990).

Due to this market thickness effect, wages are correlated with language group populations. First, the wages in the majority labor market are increasing as the majority population increases. However, the wages in the minority labor market are decreasing as the majority population increases. The former is a direct result of the market thickness effect. The latter is because more minority workers will enter the majority labor market as the wages in the majority labor market increase. The minority labor market retains less workers, and hence offers lower wages.

Second, the wages in the minority labor market are increasing in the minority population due to the market thickness effect. However, the wages in the majority labor market can either increase or decrease when the minority population increases. The latter is because the majority labor market can have either more or less minority workers if the minority population increases. On the one hand, the additional minority workers may enter the majority labor market. On the other hand, higher wages in the minority labor market may attract minority workers who previously work in the majority labor market to the minority market.

I test the above implications of the model using the 2001 Canadian

[2]See Lang (1986) or Wardhaugh (2005) for a discussion of the tacit nature of language.

Census Public Use Microdata on Individuals. One special feature of the data is the reported work language. This feature allows me to identify a worker's labor market segment by her work language. Her language group can then be identified by her home language.

I find that workers who speak English (or French in Quebec) both at work and at home get higher wages than those who speak English (or French in Quebec) at work but speak minority languages at home. However, this gap is significantly reduced after I control workers' immigration ages, and occupation fixed effects. In addition, the return to education and the return to work experience are higher for the former group of workers. The second result is robust against including workers' immigration ages and occupations.

On the other hand, minority workers who speak English (or French in Quebec) at work earn higher wages than those who speak their home languages at work. Moreover, the return to education and the return to work experience are higher for the former. Both results are robust against including workers' immigration ages and occupations.

Next, I tested wages' comparative statics with respect to population measures. However, both minority and majority populations may be endogenous. For example, a positive wage shock in a city may attract even more workers to this city, resulting in a positive correlation between language group populations and the error term of a wage equation.

To solve this problem, I constructed instrumental variables for both the minority population and majority population.[3] I employed Two Stage Least Squares (2SLS) to test the predicted effects of population measures on workers' wages. The signs of changes predicted by the theory were confirmed. This pattern of comparative statics is unique to the language theory of labor market segmentation, and hence it differentiates the theory from the human capital view of language skills held by, for example, Chiswick and Miller (1992, 1995) and Bleakley and Chin (2004), to name just a few.

This chapter also differs from the seminal work of Lang (1986), who also explores the relationship between language and wage determination. However, his economic mechanism is quite different. In his model, it is the different capital to labor ratios of different language groups that drive the wage gap. In this chapter, the basic driving force is the market thickness effect.

My model is also different from the traditional labor market segmentation literature(see Dickens and Lang, 1992, for a review of this literature). Most empirical research, such as Dickens and Katz (1987) and Dickens

[3]For details about these instruments section 2.4.4 and Appendix A.3.

and Lang (1985), focuses on inter-industry or inter-employer comparisons of wages. Language plays no role in these papers.

Various immigration policies can be analyzed within the framework of the model. For example, the policy to encourage new immigrants to relocate into small cities is likely to fail, especially for those who have lower language skills. They are much better off to stay in big cities. The model, by relating immigration policies to changes in the underlying model parameters, may be a useful tool for policy makers.

This chapter consists of three parts. In the following section, a simple model with two language groups and two labor markets is presented. Section 2.3 describes the major source of data - the 2001 Canadian Census Public Use Microdata on Individuals. Section 2.4 lays out the empirical model used to test the labor market segmentation theory. After that, Section 2.5 presents my empirical results about wage gaps and the comparative statics. Our results are then checked for their robustness in Section 2.6. Finally, Section 2.7 concludes with a discussion of the contributions of this chapter and the direction of future work.

2.2 Model

2.2.1 Workers

In a city, workers belong to two language groups (indexed by superscript k). One is the majority that has a larger population. The other has a smaller population and, hence, is the minority group.[4] The total number of workers of a language group is fixed. I denote it by n^k, where $k \in \{d, m\}$.

The labor market in each city is segmented into two submarkets (indexed by subscript l) in which the communication methods are different. In one labor market segment, the common language of business is the majority group's language. In the other labor market, the common language of business is the minority language. The number of workers in labor market l that come from language group k is denoted n_l^k, where $k \in \{d, m\}$ and $l \in \{l_d, l_m\}$.

There is no unemployment. Each worker has to sell one unit of labor. She works in only one firm. She is perfectly mobile across the two labor submarkets. The problem for her is to choose the labor submarket that

[4] The majority language in the country may not always correspond to the majority language of a particular city. One example would be the French in Quebec. While French people make up the majority in Quebec, French is not the majority language group in Canada.

maximizes her net wage. She is risk-neutral. The uncertain net wage of a worker (indexed by second subscript i) is

$$U^k_{l,i} = W^k_{l,i} - \theta^k_{l,i}, \tag{2.1}$$

where $W^k_{l,i}$ is the uncertain wage of worker i who belongs to language group k and works in labor submarket l, and $\theta^k_{l,i}$ is the cost to learn a second language. The wage is uncertain because the worker does not know how well her skills and experience will be matched with her job, i.e., how high the matching quality will be.

Workers' wages are determined by their characteristics. The first characteristic is denoted as x_i. This represents the job a worker is best suited for. Its distance from the job requirement in turn, determines the matching quality between the worker and the job. The second characteristic is denoted as α^k_i. It determines a worker's productivity relative to others. In other words, those with higher α's are more productive.

The language learning cost $\theta^k_{l,i}$ is described as

$$\theta^k_{l,i} = \begin{cases} \theta^k_l + \psi_i & \text{if } k \neq l \\ 0 & \text{if } k = l \end{cases}, \tag{2.2}$$

where θ^k_l is the average cost for a person from group k to learn the language used in labor market l and ψ_i is person i's individual cost for language learning. Notice that the choice of labor market is simultaneous to the decision to learn a new language. The language learning costs may be pecuniary or non-monetary. An example of a pecuniary cost is the time spent in a language training program and the tuition. The non-monertary cost may include lost social networks and changes in lifestyle.

Before choosing the labor market segment or the work language, workers know their own characteristics. However, they do not know the matching firm or the job they will get, *ex ante*. They form expectations about the quality of the match based on their beliefs about the number of firms in each labor market.

2.2.2 Firms

Firms decide whether to enter either of the two labor submarkets. There are no barriers to entry. The number of firms in each labor submarket is denoted m_l, where $l = \{l_d, l_m\}$. Firms produce a single product with infinitely elastic demand. Therefore, the price of the product is a constant, and the price can be normalized to one.

Firms differ in terms of their job requirements, as described by the address y on a unit circle. Correspondingly, the worker's first characteristic x_i can be modeled also as a location on a unit circle. This matching mechanism closely follows that presented by Helsley and Strange (1990). Let parameter β_l^k be the lost productivity due to unit distance of mismatch between a worker and a firm. The output of a match (x_i, y) is

$$\alpha_i^k - \beta_l^k |x_i - y| = \alpha^k + \xi_i - \beta_l^k |x_i - y|. \tag{2.3}$$

Note that α_i^k is the output of worker i if her x_i exactly matches the firm's job requirement y. It is decomposed into α^k and ξ_i: the former is the average output of a worker from group k and the latter is individual heterogeneity in productivity. $|x_i - y|$ is the mismatch between the worker and the firm.

The parameter β_l^k captures the effect of language on matching quality. Language skills matter regarding the cost of a mismatch to the worker's productivity. A mismatched worker can be more productive if she has better language skills because she can easily learn from her coworkers how to carry out a new task. This implies that $\beta_{l_d}^m > \beta_{l_d}^d$, saying that the cost of a mismatch is higher for minority workers in the majority labor market. It is also assumed that $\beta_{l_m}^d > \beta_{l_m}^m$, meaning that the cost of a mismatch is higher for majority group workers in the minority labor market.

In addition, I assume that $\beta_{l_d}^m > \beta_{l_m}^m$, meaning that a unit distance of mismatch is more costly to a minority worker if she works in the majority labor market. Analogously $\beta_{l_m}^d > \beta_{l_d}^d$, prescribing that a mismatch is more costly for a majority group worker if she works in the minority labor market.

Denote the set of majority workers hired by a firm in labor market l and with characteristic y as $\Omega_l^d(y)$. Let $\Omega_l^m(y)$ denote the set of minority workers hired by this firm. I also let $\Omega_l^d(y)$ and $\Omega_l^m(y)$ represent the number of workers within the respective sets. Therefore, this firm's total output (or the total revenue because the price of the output is normalized to one) can be written as

$$\begin{aligned} q_l(y, \Omega_l^d(y), \Omega_l^m(y)) \;=\; & \sum_{i \in \Omega_l^d(y)} (\alpha^d + \xi_i) \;-\; \beta_l^d \sum_{i \in \Omega_l^d(y)} |x_i - y| \\ & + \sum_{i \in \Omega_l^m(y)} (\alpha^m + \xi_i) \;-\; \beta_l^m \sum_{i \in \Omega_l^m(y)} |x_i - y|. \end{aligned} \tag{2.4}$$

The cost of producing these goods is

$$\kappa_l\Big(y, \Omega_l^d(y), \Omega_l^m(y), C_l\Big) \;=\; C_l + \sum_{i \in \Omega_l^d(y)} W_{l,i}^d + \sum_{i \in \Omega_l^m(y)} W_{l,i}^m. \tag{2.5}$$

C_l is the fixed cost of setting up a new firm in labor submarket l, and $W^k_{l,i}$ is the wage paid to a worker i who belongs to language group $k = \{d, m\}$. In the next subsection, a wage bargaining process is specified that determines the wages.

Firms know the distribution of workers' productivity and of workers' language learning costs. They also know that each worker's characteristic x_i is uniformly distributed on the unit circle. Moreover, they know the number and addresses of firms in each labor market. Based on this knowledge, firms form expectations about their profits and make their entry decisions.

2.2.3 Wage Bargaining

The wage of a worker is negotiated between a firm and a worker. If two parties have equal bargaining powers, the outcome is an equal split of the surplus. The surplus from a match is the output of a worker. The wage of worker i who is matched with a firm with location y on the unit circle in labor market l is hence

$$W^k_{l,i} = \frac{1}{2}[\alpha^k + \xi_i - \beta^k_l |x_i - y|], \tag{2.6}$$

where $k \in \{d, m\}$ and $l \in \{l_d, l_m\}$. Here I implicitly assume that firms observe the worker's individual characteristics ξ_i once the negotiation begins.[5] The combination of equation (2.6), (2.1), and (2.2) fully describes a worker's objective when deciding which labor market segment to work in. Similarly, equation (2.6), (2.4), and (2.5) describe the firm's objective.

2.2.4 Rational Expectations Equilibrium

Each firm expects to hire workers whose characteristics on the unit circle are closest to its own. On the other hand, each worker also expects to be employed by a firm whose location on the unit circle is closest to her skill. In fact, these expectations are rational in that each worker maximizes her net wage and each firm maximizes its profits. The resulting equilibrium is hence a 'rational expectations equilibrium'.

By symmetry, a firm with characteristic y has a market area $(y - \frac{1}{2m_l}, y + \frac{1}{2m_l})$. Each firm is assumed to meet all the workers. Whether the worker

[5] Note here it is not necessary for both parties to have zero outside options in order to get an equal split of the surplus. In light of the "Outside Options Principle" (Sutton, 1986), what I really need is that both parties' outside options cannot be higher than what they can get from this bargaining game. This turns out to be true in equilibrium. See the next section for details.

is employed by the firm is a Bernoulli random variable. Since the market area of a firm is $1/m_l$, the probability of success is $1/m_l$. If the repeated matching experiment is independent across trials, the number of majority group workers employed by the firm $\Omega_l^d(y)$ follows a binomial distribution with parameters n_l^d and $1/m_l$.[6] Therefore, the expectation of $\Omega_l^d(y)$ can be expressed as $E[\Omega_l^d(y)] = n_l^d/m_l$. Similarly, the expected employment of minority workers is $E[\Omega_l^m(y)] = n_l^m/m_l$. Note that neither depends on y.

The expected matching quality does not depend on language group because the distribution of a worker's skill x_i is independent of her language group. The expected quality of the match can be expressed as $E[|x-y| : x_i \in (y-\frac{1}{2m_l}, y+\frac{1}{2m_l})] = 1/4m_l$. It can be confirmed that the expected matching quality increases with m_l, since a smaller $1/4m_l$ indicates a higher quality match.

2.2.5 Expected Net Wage and Profit

A worker maximizes her expected net wage. Given equation (2.6), (2.1), and (2.2), the expected net wage of a worker who has characteristics $(\alpha_i^k, \theta_{l,i}^k, x_i)$ is

$$E\left(U_{l,i}^k\right) = \begin{cases} \frac{1}{2}\left[\alpha^k + \xi_i - \frac{\beta_l^k}{4m_l}\right] - \theta_l^k - \psi_i & \text{if } l \neq k, \text{ and } k \in \{d,m\},\ l \in \{l_d, l_m\} \\ \\ \frac{1}{2}\left[\alpha^k + \xi_i - \frac{\beta_{l_k}^k}{4m_{l_k}}\right] & \text{if } l = k, \text{ and } k \in \{d,m\},\ l \in \{l_d, l_m\} \end{cases} . \quad (2.7)$$

Note that $\alpha_i^k = \alpha^k + \xi_i$ and $\theta_{l,i}^k = \theta_l^k + \psi_i$.

It is apparent that a minority worker chooses the majority labor market if and only if $\frac{\beta_{l_m}^m}{8m_{l_m}} - \frac{\beta_{l_d}^m}{8m_{l_d}} - \theta_{l_d}^m - \psi_i > 0$. She can get potentially higher matching quality in the majority market, represented by $\frac{\beta_{l_m}^m}{8m_{l_m}} - \frac{\beta_{l_d}^m}{8m_{l_d}}$. It is positive if m_{l_d} is larger than m_{l_m} and if $\beta_{l_d}^m$ is not much larger than $\beta_{l_m}^m$. On the other hand, she has to pay a cost to learn the majority language, represented by $\theta_{l_d,i}^m$.

However, most majority workers do not face such a tradeoff. First, the matching quality term $\frac{\beta_{l_d}^d}{8m_{l_d}} - \frac{\beta_{l_m}^d}{8m_{l_m}}$ is always negative. Second, the language learning cost $\theta_{l_m,i}^d$ is positive except in extreme cases. Therefore, it is highly unlikely that a majority worker would switch labor markets. The allocation of minority workers is summarized in the following lemma assuming that majority workers stay in the majority labor market.

[6] See Appendix A.1.1 for the detailed derivation of this and the following paragraph.

Lemma 2.1. *If (1) majority workers only work in the majority labor market, and (2) $\psi_i \sim N(0, \sigma^2)$, then the following results hold.*

1. $n^m_{l_d}$ *is a binomial random variable with a mean of* $n^m \cdot \Phi(\mu/\sigma)$*, where* $\mu = \frac{\beta^m_{l_m}}{8m_{l_m}} - \frac{\beta^m_{l_d}}{8m_{l_d}} - \theta^m_{l_d}$ *and* $\Phi(.)$ *is the standard normal cumulative distribution function.*

2. $n^m_{l_m}$ *is a binomial random variable with an expectation of* $n^m \cdot (1 - \Phi(\mu/\sigma))$.

Proof. A minority worker chooses the majority labor market if and only if $\psi_i < \frac{\beta^m_{l_m}}{8m_{l_m}} - \frac{\beta^m_{l_d}}{8m_{l_d}} - \theta^m_{l_d}$. Since $\psi_i \sim N(0, \sigma^2)$, her probability of entering the majority labor market is $\Phi(\mu/\sigma)$. Because ψ_i's are independent across workers, $n^m_{l_d}$ is a binomial random variable with an expectation of $n^m \cdot \Phi(\mu/\sigma)$. Analogously, we can get the result for $n^m_{l_m}$. □

Assumption 2.1.
1. Majority workers only work in the majority labor market;

2. ξ_i and ψ_i are multivariate normal with an expectation of $(0, 0)^T$ and a variance-covariance matrix $\begin{pmatrix} s^2 & \rho s\sigma \\ \rho s\sigma & \sigma^2 \end{pmatrix}$;

3. x_i is independent of both ξ_i and ψ_i;

4. ξ_i, ψ_i, and x_i are identically and independently distributed across all workers.

The first assumption is realistic based on Canadian Data. In cities outside Quebec, about 97% of English-speaking workers speak English at work, another 2% speak French at work, and the remaining 1% speak other languages at work. Since the majority language outside Quebec is English, the data shows that majority workers are very unlikely to work in the minority labor market. The second assumption specifies the joint distribution of ψ_i and ξ_i. The correlation coefficient ρ captures the possibility that high ability workers may be more effective in both production and language learning. This is related to the selection issues in the empirical part. The last two assumptions are mainly technical, but they are not particularly unrealistic.

Lemma 2.2. *Under Assumption 2.1, the following statements hold. The expected profit of a firm in the majority labor market is*

$$E(\pi_{l_d}) = \frac{n^d}{2m_{l_d}}\left(\alpha^d - \frac{\beta^d_{l_d}}{4m_{l_d}}\right) + \frac{n^m \cdot \Phi(\mu/\sigma)}{2m_{l_d}}\left[\alpha^m - \rho s\frac{\phi(\mu/\sigma)}{\Phi(\mu/\sigma)} - \frac{\beta^m_{l_d}}{4m_{l_d}}\right] - C_{l_d}. \quad (2.8)$$

The expected profit of a firm in the minority labor market is

$$E(\pi_{l_m}) = \frac{n^m \cdot (1-\Phi(\mu/\sigma))}{2m_{l_m}}\big[\alpha^m + \rho s \frac{\phi(\mu/\sigma)}{1-\Phi(\mu/\sigma)} - \frac{\beta^m_{l_m}}{4m_{l_m}}\big] - C_{l_m}. \tag{2.9}$$

Note $\mu = \frac{\beta^m_{l_m}}{8m_{l_m}} - \frac{\beta^m_{l_d}}{8m_{l_d}} - \theta^m_{l_d}$.

Proof. The proof is contained in Appendix A.1.2. □

The expected profit in the majority labor market $E(\pi_{l_d})$ includes three components: the net revenue from majority workers, the net revenue from minority workers, and the fixed cost of operating a firm. Notice the additional component $-\rho s\frac{\phi(\mu/\sigma)}{\Phi(\mu/\sigma)}$ that represents an adjustment due to the selection of high ability minority workers into the majority labor market.[7] A similar interpretation applies to the expected profit in the minority labor market.

The necessary condition for equilibrium are that firms in both labor markets earn zero profits. In fact, the two zero profits conditions completely characterize the number of firms. Other variables of interest can then be expressed as a function of the number of firms. I should note that the minority labor market may not exist in equilibrium. In addition, multiple equilibria where both labor market exist are also possible. Any comparative statics necessarily hinge upon restrictive assumptions.

Proposition 2.1. *If (1) conditions in Assumption 2.1 hold, (2) the model parameters are such that there is at least one pair of* (m_{l_d}, m_{l_m})*, where* $m_{l_d} > 0$ *and* $m_{l_m} > 0$*, (3)* $\rho = 0$*, and (4) the Jacobian* J *of the zero profit conditions evaluated at the equilibrium pair* (m_{l_d}, m_{l_m}) *is negative definite, then the following comparative statics hold for the equilibrium pair* (m_{l_d}, m_{l_m})*.*

[7] Suppose high ability workers tend to incur less costs in learning a language and to have higher productivity. This means ψ_i and ξ_i are negatively correlated or $\rho < 0$. This in turn implies that the adjustment term is positive, meaning the firm currently earns more net revenue from minority workers than the case when there is no selection.

	m_{l_d}	m_{l_m}
Majority Population (n^d)	+	−
Minority Population (n^m)	?	?
Language Learning Cost ($\theta^m_{l_d}$)	−	+
Communication Cost #1 ($\beta^d_{l_d}$)	−	+
Communication Cost #2 ($\beta^m_{l_d}$)	−	+
Communication Cost #3 ($\beta^m_{l_m}$)	+	−
Firm Operating Cost #1 (C_{l_d})	−	+
Firm Operating Cost #2 (C_{l_m})	+	−
Majority Productivity (α^d)	+	−
Minority Productivity (α^m)	?	?

Proof. See Appendix A.1.3 for detailed derivation. □

The first assumption is familiar. The second states that there are two submarkets in equilibrium. This is the case when it is appropriate to discuss the comparative statics. The third assumption is added to simplify the analysis. I will consider non-zero ρ in the following numerical example. The last assumption selects an equilibrium in which the numbers of firms and the equilibrium profits change in the opposite direction. This equilibrium is stable in the sense that small shocks to profits do not lead to divergence from the equilibrium.[8]

The basic parameter setup of the numerical example is as follows: $n^d = 10000$, $n^m = 5000$, $\alpha^d = \alpha^m = 10000$, $\beta^d_{l_d} = \beta^m_{l_m} = 15000$, $\beta^m_{l_d} = 18000$, $\rho = -0.9$, $s = 5000$, $\sigma = 10$, $\theta^m_{l_d} = 20$, and $C_{l_d} = C_{l_m} = 100000$. Note that I allow $\rho \neq 0$ in this numerical exercise.

Figure 2.2 to Figure 2.4 show the pairs of (m_{l_d}, m_{l_m}) that satisfy the two zero profit conditions. The thick solid line represents the zero profit condition of the majority market. The thin solid line corresponds to the zero profit condition of the minority market. The intersection of the two is therefore an equilibrium.[9]

The left graph of Figure 2.2 shows the comparative statics with respect to majority population. As the majority population increases, the zero profit curve of the majority market shifts outward, while that of the minority market remains unchanged. As a result, the number of majority firms increases,

[8] Suppose firms enter if there are positive profits and exit if the profits are negative. A necessary and sufficient condition for the stability of the equilibrium is that (1) $J(1,1) + J(2,2) < 0$ and (2)$|J| > 0$. These two conditions are implied by assumption (4) of Proposition 2.1.

[9] There are in fact multiple solutions. However, the one shown in this figure is the only stable solution.

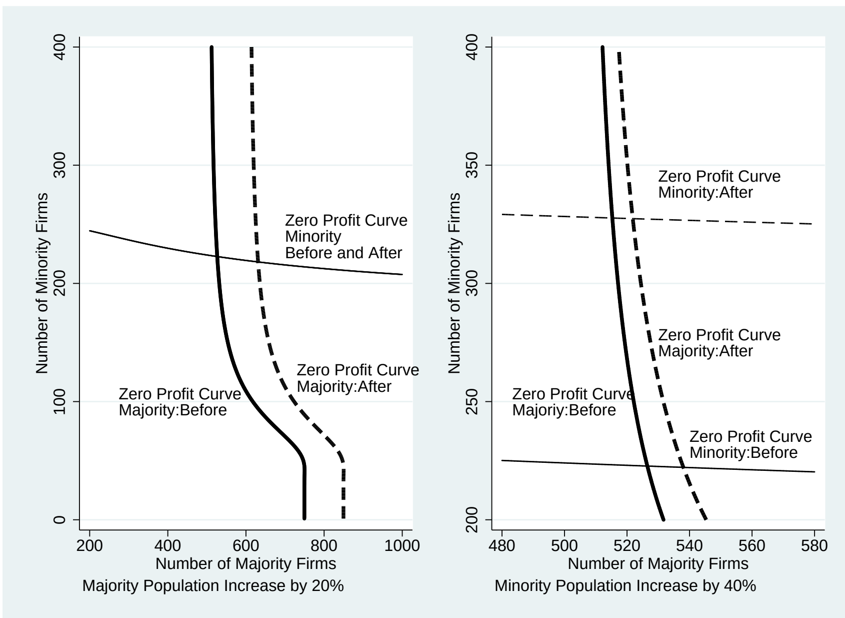

Figure 2.2: Number of Firms and the Language Group Populations

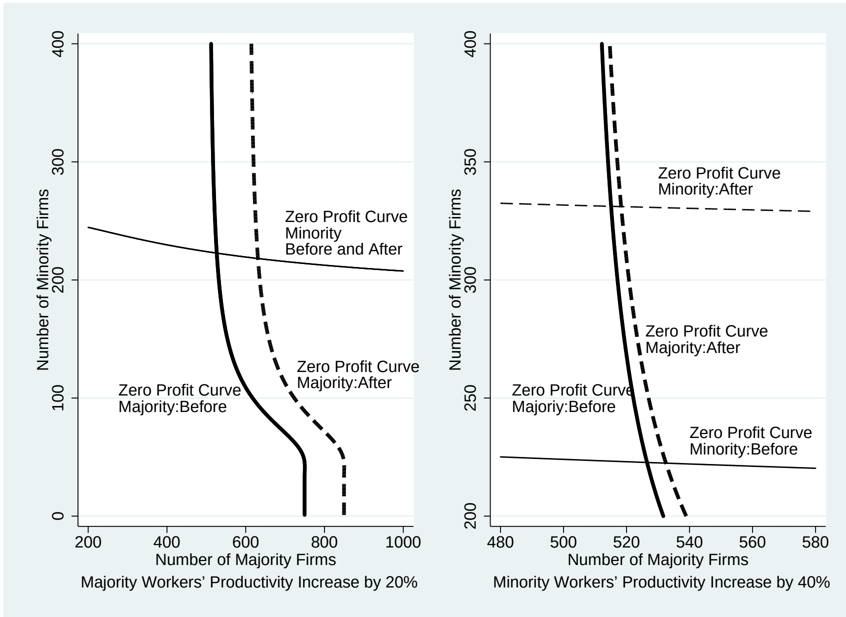

Figure 2.3: Number of Firms and the Average Productivities

while that of minority firms decreases. This happens because the matching quality in the majority labor market improves as majority population increases. More firms enter this labor market. At the same time, more minority workers enter the majority market in search of higher wages. Profits in in the minority labor market decrease. Firms exit the minority labor market.

The second graph of Figure 2.2 shows the comparative statics with respect to the minority population. Both zero profit curves shift outward. The number of minority firms increases significantly, while that of majority firms decreases slightly. Although Proposition 2.1 does not have a definite answer about it, this example indicates that the number of minority firms is more sensitive to minority population.

Figure 2.3 illustrates the comparative statics with respect to average productivity of the two groups, namely α^d and α^m. The results are similar to those in Figure 2.2. The economic mechanism is also similar. Again, the number of minority firms seems to be more sensitive to changes in the average productivity of α^m.

Figure 2.4 shows the comparative statics with respect to $\theta^m_{l_d}$ and $\beta^m_{l_d}$.

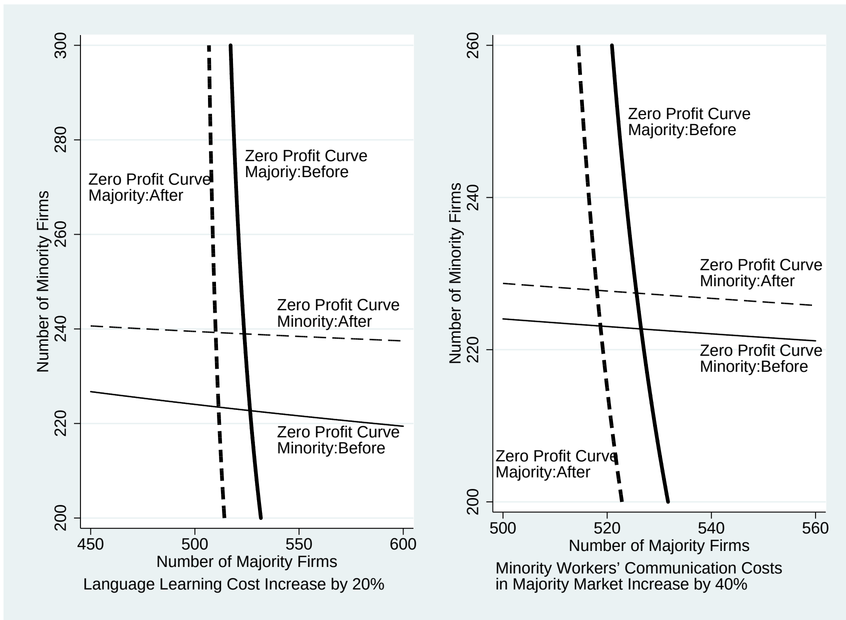

Figure 2.4: Number of Firms and the Language Costs

As these language related costs increase, the number of majority firms decrease and the number of minority firms increase. The increase in those costs discourages minority workers from entering the majority market. The matching quality and potential profit in the minority market increase, so more firms enter the minority market. On the other hand, the profit in the majority market diminishes, and firms exit from the majority market.

Comparative statics with respect to other parameters follow the same line of thought. It is also straightforward but tedious to generalize the two-ethnic-group, two-labor-submarket framework into a multi-ethnic-group and multi-labor-submarket model. Appendix A.2 contains a sketch of such a generalization. The comparative statics become messier but the message is still the same.

2.2.6 Empirical Implications

In this section, I will discuss the empirical implications of the theory. The first key implication is summarized in the following proposition.

Proposition 2.2. *There are two types of wage gaps: the Within-Labor-Market Wage Gap and the Within-Language-Group Wage Gap. Under Assumption 2.1, the Within-Labor-Market Wage Gap can be expressed as*

$$E(W^d_{l_d} - W^m_{l_d}) = \frac{1}{2}[\alpha^d - \alpha^m + \frac{\beta^m_{l_d}}{4m_{l_d}} - \frac{\beta^d_{l_d}}{4m_{l_d}}]. \tag{2.10}$$

Under Assumption 2.1, the Within-Language-Group Wage Gap can be expressed as

$$E(W^m_{l_d} - W^m_{l_m}) = \frac{1}{2}[\frac{\beta^m_{l_m}}{4m_{l_m}} - \frac{\beta^m_{l_d}}{4m_{l_d}}]. \tag{2.11}$$

Proof. These two expressions are directly implied by equation (2.6) and assumption 2.1. □

Note that $W^d_{l_d}$ is the wage earned by a majority group worker. The wage of a minority worker who works in the majority labor market is denoted as $W^m_{l_d}$. The wage of a minority worker in the minority labor market is denoted as $W^m_{l_m}$. I will test the existence of the two wage gaps empirically.

The second key implication is the comparative statics of the wages with respect to the underlying model parameters. The following proposition summarizes the result.

Proposition 2.3. *Under assumptions in Proposition 2.1, the following comparative statics hold for $E(W^d_{l_d})$, $E(W^m_{l_d})$, and $E(W^m_{l_m})$:*

	$E(W^d_{l_d})$	$E(W^m_{l_d})$	$E(W^m_{l_m})$
Majority Population (n^d)	+	+	−
Minority Population (n^m)	?	?	?
Language Learning Cost ($\theta^m_{l_d}$)	−	−	+
Communication Cost #1 ($\beta^d_{l_d}$)	−	−	+
Communication Cost #2 ($\beta^m_{l_d}$)	−	−	+
Communication Cost #3 ($\beta^m_{l_m}$)	+	+	−
Firm Operating Cost #1 (C_{l_d})	−	−	+
Firm Operating Cost #2 (C_{l_m})	+	+	−
Majority Productivity (α^d)	+	+	−
Minority Productivity (α^m)	?	?	?

Proof. The derivation is straightforward. Differentiate $E(W^d_{l_d,i})$, $E(W^m_{l_d,i})$, and $E(W^m_{l_m,i})$ with respect to the population measures n^d and n^m and use the results in Proposition 2.1. □

Notice the link between Proposition 2.3 and Proposition 2.1. It is easy to confirm that both $E(W_{l_d}^d)$ and $E(W_{l_d}^m)$ move positively with m_{l_d} and that $E(W_{l_m}^m)$ moves positively with m_{l_m}. Therefore, Figure 2.2 to Figure 2.4 provide visual help to understand the results of Proposition 2.3.

One may notice the absence of the supply effect on wages. Indeed, this comes from the assumption of a fixed output price. If the output demand is downward sloping, the labor demand curve is likely to be downward sloping, so the supply effect matters. In the model, I focus on the market thickness effect which leads to a upward sloping labor demand curve. Whether the supply effect is more important than the market thickness effect is a subject of our empirical analysis.

The results in Proposition 2.3 are interesting because the effects of many immigration policies can be traced back to changes in the parameters of the model. I have classified the underlying parameters into several categories: the populations of the two language groups, the language-related costs, the fixed costs of starting a business in different labor markets, and the productivity of workers from different language groups. The framework in this chapter is hence a good tool to analyze the consequences of various immigration policies.

It is suggested that the government should encourage new immigrants to settle in small cities where they can better mingle into the society (See CIC, 2001). This decreases the minority population in big cities, which in turn reduces wages in the minority labor market and increases wages in the majority labor market. At least in the short run, the workers in the ethnic enclaves will bear the cost. The workers in the majority labor market may be better off or worse off (See Figure 2.2).

It has also been widely discussed that the government should facilitate the recognition of overseas qualifications earned by new immigrants. In fact, the Canadian federal government's Foreign Credential Recognition program does exactly this. This policy is equivalent to increasing perceived α^m by employers. Minority workers in ethnic enclaves benefit from such a policy, while worker in the majority market can either benefit or suffer from such a policy (See Figure 2.3). Policies that reduce workplace discrimination have the same effects.

Another prominent policy is to subsidize language learning, for example setting up free language lessons or subsidizing community activities. The Language Instruction for Newcomers to Canada by the Citizenship and Immigration Canada is one example. This policy reduces $\theta_{l_d}^m$ and $\beta_{l_d}^m$. Again, those who remain in the ethnic enclaves are worse off, while those who leave the ethnic enclaves and the majority workers are better off (See Figure 2.4).

The government can also implement programs that help immigrants to start small businesses. They reduce the cost of setting up a firm. The effects depend on whether those entrepreneurs set up businesses in the majority market or the minority market. If the latter is the case, workers in the ethnic enclave will benefit, while those outside the enclave will suffer.

I have shown that the language theory of labor market segmentation has rich implications. The test of the theory is however limited due to data availability. In the dataset I use, there are good measures for populations n^d and n^m. There are also proxies for $\theta^m_{l_d}$ and $\beta^m_{l_d}$. The language distance measure proposed by Chiswick and Miller (2004) is a good candidate. Age at immigration is another.[10] Unfortunately, some parameters are difficult to measure. These include the costs of operation for a firm and the within-group communication costs. Therefore, we focus on the existence of the two types of wage gaps and the comparative statics with respect to populations.

2.3 Data

The main source of data is the 2001 Census of Canada Public Use Microdata File (henceforth 2001 Canada PUMF or PUMF) on individuals. This dataset is a 2.7% sample of the Canadian population. Because this chapter is only interested in urban workers, I exclude all individuals who do not reside in a Census Metropolitan Area (CMA) from our analysis. This exclusion reduces the number of observations to 496,611, representing a population of 18,348,790. There are 19 Census Metropolitan Areas in the data.

I identify a person's language group by her language at home.[11] There are 14 language groups: English, French, Aboriginal, German, Italian, Spanish, Portuguese, Polish, Chinese, Austro-Asiatic, Arabic, Punjabi, other Indo-Iranian, and other.

Table 2.1 reports the detailed population distribution over the language groups of selected metropolitan areas. I chose Montréal, Toronto, and Vancouver because they have the largest populations of minority workers. People who speak European languages live predominantly in the eastern part of Canada. For example, the Portuguese speaking population in Toronto is about 71,847, while that in Vancouver is 3,398. Punjabi and Chinese people, on the other hand, are more likely to stay in Vancouver. About an equal number of these people live in Toronto and Vancouver. Another fact shown

[10] We also need a variable to differentiate between $\theta^m_{l_d}$ and $\beta^m_{l_d}$. This is called an 'exclusion restriction'. See Section 2.4.1 for details.

[11] Details are in Appendix A.3.

Table 2.1: Language at Home Distribution (Selected CMAs)

Language spoken at home	Canada CMAs	Montréal	Toronto	Vancouver
English	11,651,353	548,315	3,316,832	1,436,313
French	3,762,355	2,386,757	27,783	10,685
Aboriginal	4,520	370	296	296
German	53,929	5,054	14,442	6,534
Italian	216,159	66,711	100,493	7,271
Spanish	173,543	52,557	65,051	15,016
Portuguese	118,893	17,480	71,847	3,398
Polish	122,485	11,009	59,540	7,357
Chinese	708,504	38,567	305,652	253,623
Austro-Asiatic	116,179	25,402	35,110	17,825
Arabic	159,728	60,386	42,765	4,401
Punjabi	202,186	8,654	85,326	71,802
Other Indo-Iranian	284,211	30,514	161,103	40,703
Other	773,671	128,829	361,604	92,252

by this table is that these three cities account for around two thirds of the populations of most minority groups.

I identify a worker's labor market segment by her work language. The reported work language is a special feature of this dataset.[12] In this sense, this data provides a unique testing ground for the language theory of labor market segmentation. There are 17 work language groups: English, French, Aboriginal, German, Netherlandic, Italian, Spanish, Portuguese, Polish, Ukrainian, Greek, Chinese, Austro-Asiatic, Arabic, Punjabi, other Indo-Iranian, and other. Notice the three additional language groups: Netherlandic, Ukrainian, and Greek.

Table 2.2 shows the population distribution of work language groups. The asymmetric allocation of minority workers is even more conspicuous than that shown in Table 2.1. There seems to be both a Polish and a Ukrainian labor submarket in Toronto, while there seem to be no such labor submarkets in Vancouver. I should note that the number of observations on which these population estimates are based is about 1/37 of that reported in the table.

[12] See Appendix A.3 for details.

Table 2.2: Language at Work Distribution (Selected CMAs)

Language spoken at work	Canada CMAs	Montréal	Toronto	Vancouver
English	8,243,865	367,393	2,613,355	107,043,9
French	2,222,645	1,498,024	16,399	3,810
Aboriginal	996	74	0	0
German	5,497	185	961	850
Netherlandic	590	37	37	74
Italian	16,258	6,652	7,281	332
Spanish	16,570	4,958	6,399	1,480
Portuguese	15,118	1,590	10,761	185
Polish	6,539	443	3,772	370
Ukrainian	4,058	185	2,440	222
Greek	6,137	4,030	1,702	0
Chinese	111,633	5,067	46,555	47,403
Austro-Asiatic	11,524	1,886	3,218	3,290
Arabic	7,436	3,957	1,406	258
Punjabi	13,598	444	4,070	7,431
Other Indo-Iranian	16,895	1,221	8,360	4,842
Other	35,804	4,107	15,202	7,802

2.4 Empirical Strategy

2.4.1 Models of Wages and Labor Market Choices

In the model, wages are complicated non-linear functions of both individual characteristics and language group characteristics. For this reason, I adopt a log-linear specification of wages.

$$\ln W^d_{l_d,i} = X_i\gamma^d + G\delta^d + \epsilon^d_i, \tag{2.12}$$

$$\ln W^m_{l_d,i} = X_i\gamma^m_{l_d} + G\delta^m_{l_d} + \epsilon^m_{l_d,i}, \tag{2.13}$$

$$\ln W^m_{l_m,i} = X_i\gamma^m_{l_m} + G\delta^m_{l_m} + \epsilon^m_{l_m,i}. \tag{2.14}$$

Log wages are measured as the logarithm of workers' hourly earnings. Notice the subscript and superscript associated with model coefficients and error terms.

X_i is a vector of individual characteristics: education, experience, experience squared, sex, marital status, interaction of sex and marital status, a dummy indicating whether a person immigrated after age 19, and occupation fixed effects. The dummy indicating whether a person immigrated after age 19 is included to control for the adaptation of immigrants. It is also a proxy for language skills. Occupation dummies are included to differentiate this theory from a theory in which minority workers and majority workers sort into different occupations (Card, 2001).

G is a vector of group-wide variables, which includes the majority population in the city in which the worker lives, the population of the worker's language group in the city, and possible other variables. These variables are omitted in testing the two wage gaps. They are included in testing the comparative statics.

For a minority worker, the choice of labor market is based on a comparison of net wages. Similar to equation (2.7), I define the net wage gain from entering the majority labor market as

$$I_i = \ln W^m_{l_d,i} - \ln W^m_{l_m,i} - \theta^m_{l_d,i} - \psi_i = X_i\lambda_1 + Z_i\lambda_2 + G\lambda_3 + e_i. \qquad (2.15)$$

Note the additional vector Z_i, the exclusion restriction I impose on the wage equation. The error term e_i is a normal random variable.

One possible variable for the exclusion restriction is a measure of distance between a minority language and English proposed by Chiswick and Miller (2004). They use the scores of English as a Second Language students as the measure of distance between English and other languages. The score is therefore measured at language group level.

Recall that the theory says that all the workers have to acquire a formal knowledge of the majority language before they can work in that market. In addition, people who master the tacit aspects of the majority language differently earn different wages in the majority labor market. One may argue that this score does not enter the wage equation because it does not measure the tacit aspects of language learning. However, this variable may be correlated with the difficulty in learning tacit aspects of the language learning. Therefore, it is not a valid exclusion restriction.

I instead use the interaction between the language score and a worker's age at immigration dummy as the exclusion restriction. Suppose there are four immigrants. Two come from Ireland, where exposure to English is prevalent. One of them immigrated at the age of 5 and the other at 25. The other two have the same immigration age profile, but they come from Mexico, where exposure to English is less common.

The two Irish differ in terms of their tacit knowledge of English and their social networks. The two Mexicans, however, differ not only in terms of their tacit knowledge of English and their social networks, but also in their knowledge of explicit forms of English. Since the difference within an ethnicity represents the marginal effect of immigration age, the marginal effect of immigration age should be higher for the Mexicans. Formally, we should have immigration age, language score, and the interaction between the two in a labor market choice equation.

After the labor market choice, assume we have four people who all work in the majority labor market. By definition, the two Mexicans no longer differ in terms of their knowledge of explicit forms of English. Other things equal, the wage difference between the two Irish and the wage difference between the two Mexicans should be the same. This means the interaction term does not enter the wage equation. In summary, the interaction between immigration age and language score enters the choice equation but does not enter the wage equation, and it is a valid exclusion restriction.[13]

2.4.2 Within-Labor-Market Wage Gap

Recall that the Within-Labor-Market Wage Gap is the wage difference between a majority worker and a minority worker in the majority labor market. Define $\ln W_{l_d,i}$ as $D_i \ln W^d_{l_d,i} + (1-D_i) \ln W^m_{l_d,i}$, where $D_i = 1$ indicates a majority worker and $D_i = 0$ indicates a minority worker. If $E(X_i|G) = E(X_i)$, we can omit group wide variables in equation (2.12) and (2.13) and write

$$E(\ln W_{l_d,i}|X_i, D_i) = \mu_0 + D_i \cdot \mu_1 + X_i \boldsymbol{\mu_2} + D_i(X_i - E(X))\boldsymbol{\mu_3} + (1-D_i)\frac{\phi(.)}{\Phi(.)}\mu_4. \quad (2.16)$$

Under standard conditions in the treatment effect literature, we can estimate the Within-Labor-Market Wage Gap by running the above regression over all workers in the majority labor market.[14] The estimated wage gap is μ_1 for a worker with average characteristics $E(X)$. The estimated wage gap is $\mu_1 + (X_i - E(X))\boldsymbol{\mu_3}$ for a worker with characteristic X_i, which includes education, work experience, and other individual characteristics.

The inverse Mill's ratio term $\frac{\phi(.)}{\Phi(.)}$ is included because only minority workers who choose the majority labor market are in the sample. It is constructed from equation (2.15) with group-wide variables omitted.

It is possible that the above estimated wage gap is a reflection of discrimination. To differentiate from such an alternative, I add worker's visible minority status V_i, an interaction $L_i \times V_i$, and the interaction between the status and the de-meaned individual characteristics $V_i(X_i - E(X))$ to the previous regression. $V_i = 1$ indicates that a worker is a visible minority and $V_i = 0$ indicates otherwise.

[13] I do not have similar language scores for French. However, when I model workers' labor market or work language choice in Quebec, I do need such a score. I assume that a minority individual's French score is the same as her English score.

[14] See Wooldridge (2002) for a complete discussion of the conditions underlying the use of switching regression in order to estimate Average Treatment Effect.

2.4.3 Within-Language-Group Wage Gap

Recall that Within-Language-Group Wage Gap is the wage difference between two minority workers who work in different labor markets. Define $\ln W_i^m$ as $L_i \ln W_{l_d,i}^m + (1 - L_i) \ln W_{l_m,i}^m$. L_i indicates the worker's labor market status, which equals 1 if it is the majority labor market but equals 0 otherwise. If $E(X_i|G) = E(X_i)$, we can omit group wide variables in equation (2.13) and (2.14) and write

$$\begin{aligned} E(\ln W_i^m | X_i, L_i) &= \mu_0 + L_i \mu_1 + X_i \boldsymbol{\mu_2} + L_i (X_i - E(X)) \boldsymbol{\mu_3} \\ &+ (1 - L_i) \frac{\phi(.)}{1 - \Phi(.)} \mu_4 + L_i \frac{\phi(.)}{\Phi(.)} \mu_5. \end{aligned} \tag{2.17}$$

We can estimate the Within-Language-Group Wage Gap by running the above regression over all minority workers.[15] The estimated wage gap is μ_1 for a worker with average characteristics $E(X)$. The estimated wage gap is $\mu_1 + (X_i - E(X))\boldsymbol{\mu_3}$ for a worker with characteristic X_i.

In the sample, we only observe workers who have made their labor market choices. I follow the framework of Lee (1982) in addressing this self-selection problem. The difference is that I include both selection correction terms $\frac{\phi(.)}{\Phi(.)}$ and $\frac{\phi(.)}{1-\Phi(.)}$ in one switching regression instead of running two separate regressions. The selection correction terms are both constructed from equation (2.15), with the group-wide variables omitted.

It is also possible that the above estimated wage gap is because of the different discriminative treatment of minorities across labor markets. To differentiate from such alternatives, I add the worker's visible minority status V_i and the interaction between the visible minority status and the de-meaned individual characteristics $V_i(X_i - E(X))$ to the previous regression. Moreover, the interaction between visible minority status and labor market status $L_i \times V_i$ is also added to control for different degrees of discrimination in the two labor markets.

2.4.4 Tests of Comparative Statics

In this section, I include population measures to test their marginal effects on individual wages. I again adopt the method of Lee (1982) to address the selection issues. I run two separate wage regressions for minority workers in the majority labor market and minority workers in the minority labor market, respectively.

$$E(\ln W_{l_d,i}^m | X_i, \tilde{G}, n^m, n^d, L_i = 1) \quad = \quad X_i \boldsymbol{\mu_{1d}} + \tilde{G} \boldsymbol{\mu_{2d}} + \mu_{3d} \ln n^m$$

[15] Again, see Wooldridge (2002) for the conditions needed.

$$+ \quad \mu_{4d} \ln n^d + \frac{\phi(.)}{\Phi(.)} \mu_{5d}. \tag{2.18}$$

$$\begin{aligned} E(\ln W^m_{l_m,i} | X_i, \tilde{G}, n^m, n^d, L_i = 0) &= X_i \mu_{\mathbf{1m}} + \tilde{G} \mu_{\mathbf{2m}} + \mu_{3m} \ln n^m \\ &+ \mu_{4m} \ln n^d + \frac{\phi(.)}{1 - \Phi(.)} \mu_{5m}. \end{aligned} \tag{2.19}$$

I include in X_i the individual characteristics as well as the worker's visible minority status and its interactions with de-meaned education, experience, and experience squared. Language group dummies are contained in $\tilde{G}$. The log of the majority population in a city is denoted as $\ln n^d$. The log of the worker's group population in the city is denoted as $\ln n^m$. The selection correction terms are constructed from equation (2.15), in which group-wide variables $\ln n^d$ and $\ln n^m$ are included.

According to the theory, we should observe a positive μ_{4d} and either a positive or a negative μ_{3d}. On the other hand, we should observe a negative μ_{4m} and a positive μ_{3m}.

There is a caveat in estimating the two wage regressions. Both the majority and minority populations may be endogenous. For example, a wage shock in a city may attract more workers to come to this city, resulting in a positive correlation between population and wages. This issue has been discussed by Glaeser and Mare (2001).

I employ Two Stage Least Squares (2SLS) to address this problem. I construct two instruments for minority population. The first instrument is the distance from the origin of the language group to the city where the worker lives. It is hard to imagine that the wage of a worker depends on the distance from her language of origin to where she currently lives. Table 2.1 shows that European immigrants tend to settle in eastern provinces and Asian immigrants tend to reside in western provinces. It suggests a correlation between the distance measure and language group population in a city.

The second instrument is the product of one city's 1990 share of a language group's total national population and the total national population of the group in 2000. This type of instrument is discussed by Card (1999). This predicted language group population is based on a national trend, so it is exogenous to city-specific wage shocks. On the other hand, it is correlated with the realized group population in a city in 2000.

I am less concerned about the endogeneity of the majority population. The wage shocks to minority workers would only have a limited effect on the location choices of majority workers. I use 'imputed population' as an instrument for the majority population, and this instrument is similar to the second instrument for minority population.[16]

2.4.5 Issues about the Data

In the data, there are several language groups and several potential labor markets in a city. In the theory, however, there are only two language groups and two labor markets. I extend the model to multi-group and multi-market setting in Appendix A.2. In this setting changes in one minority labor market may affect wages other minority labor markets. I call this the 'ripple effect'.

The 'ripple effect' does not matter much if there is a single group that has a much larger population than the populations of all the other groups. However, it matters if there are two groups that are similar in size: changes in the slightly smaller language group may affect other smaller groups. I suspect that Montreal and Ottawa are two examples. To alleviate this concern, I carry out the analysis on a subsample of cities, excluding cities in Quebec and the city of Ottawa.

A related problem may arise due to the bilingualism policy in Canada. When there are a lot of people who can communicate in both English and French, the boundaries between labor markets are no longer clear. In addition, the fact that there are certain advantages to becoming bilingual, such as qualifying for government jobs, complicates the analysis. Empirically, I alleviate the problem by simply excluding the English group in Quebec and the French group outside Quebec from the analysis.

2.5 Regression Results

2.5.1 Labor Market Choice

The labor market choice is modeled in equation (2.15). Table 2.3 reports the estimates. There are two subsamples: workers in all Canadian CMAs and workers in CMAs outside Quebec and Ottawa. The estimates in this table are used to construct selection correction terms in wage regressions later on.

In general, higher education leads to higher probability of selecting the majority labor market. Older people are less likely to enter the majority labor market. Those who immigrated after age 19 are less likely to choose the

[16] See Appendix A.3 for the details of those instruments.

Table 2.3: Labor Market Choices of Minority Workers

	All CMAs			CMAs outside Quebec		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Education	0.0710[a]	0.0709[a]	0.0728[a]	0.0723[a]	0.0720[a]	0.0746[a]
	(0.0064)	(0.0067)	(0.0073)	(0.0075)	(0.0081)	(0.0087)
Age	-0.0066[a]	-0.0063[a]	-0.0062[a]	-0.0060[a]	-0.0057[a]	-0.0055[a]
	(0.0015)	(0.0014)	(0.0013)	(0.0017)	(0.0015)	(0.0015)
Sex (Female=1)	-0.0046	-0.0019	0.0014	-0.0089	-0.0057	-0.0006
	(0.0199)	(0.0196)	(0.0186)	(0.0206)	(0.0203)	(0.0192)
Immigrate after 19	-0.2444[a]	-1.1562[a]	-1.1591[a]	-0.2628[a]	-1.2296[a]	-1.2426[a]
	(0.0428)	(0.2296)	(0.2684)	(0.0498)	(0.2465)	(0.2939)
Language Score (LS)	0.5323[b]	0.2527	0.1993	0.6323[b]	0.3356	0.2499[c]
	(0.2531)	(0.1911)	(0.1341)	(0.2826)	(0.2109)	(0.1460)
LS*Imm. after 19		0.4692[a]	0.4723[a]		0.5006[a]	0.5083[a]
		(0.1299)	(0.1477)		(0.1417)	(0.1628)
Log own pop.			-0.0735			-0.1064
			(0.1008)			(0.1049)
Log maj. pop.			0.0644			0.1222
			(0.1637)			(0.1705)
Intercept	-0.8529	-0.3082	-0.3361	-1.0577[c]	-0.4798	-0.8950
	(0.5213)	(0.4404)	(1.5154)	(0.5617)	(0.4744)	(1.5819)
N	49379	49379	49379	44222	44222	44222
Pseudo R^2	0.0528	0.0554	0.0580	0.0591	0.0620	0.0662

The choice variable is the worker's language at work. It equals 1 if it is English (or French in Quebec), 0 otherwise. Minority workers are those whose home languages are neither English nor French. [a] significance level of 1%. [b] significance level of 5%. [c] significance level of 10%. Standard errors are in parentheses. The error terms are clustered by individual language group within a CMA.

majority labor market. Female workers seem to be less likely to work in the majority labor market, though the difference is not statistically significant.

For all CMAs, the effect of language score has the expected positive sign. Recall that the language score is higher when the distance between two languages is smaller. The marginal effect of immigration age is lower for people who have higher language scores. This result is consistent with our discussion in Section 2.4.1, where the marginal effect of immigration age depends on whether a worker is an Irish or a Mexican. Based on Model 3, the size of the majority population in the city increases the probability of a worker choosing the majority labor market. The person's own group's population has a negative effect. The effects of the population measures are not statistically significant, however. Those results are consistent with the language theory of labor market segmentation. The results for CMAs outside Quebec are similar.

2.5.2 Within-Labor-Market Wage Gap

In this section, I show the existence of the Within-Labor-Market Wage Gap. Remember it is the wage difference between a majority worker and a minority worker in the majority labor market. The sample includes all workers who speak English at work outside Quebec or French in Quebec. I exclude workers who speak English at home in Quebec and workers who speak French at home outside Quebec. This exclusion alleviates the concern about the bilingualism policy.

Table 2.4 shows the results, first, for all CMAs and, secondly, for CMAs outside Quebec and Ottawa, respectively. The exclusion of the CMAs in Quebec and Ottawa is to address the concern about the 'ripple effect'. Since the results for the two subsamples are fairly similar, I discuss the results for all CMAs only.

Model 1 is the starting point. Since I do not include age at immigration, I am comparing one person who was born in Canada versus another person who was born in another country. I find that a majority worker, who has the average education and experience of the sample, earns 13.9% more than her minority counterpart. In addition, the former gets more for one additional year of education and work experience. The regressors, such as education, experience, and others, all have the expected signs.

Model 2 controls for age at immigration dummy and occupation fixed effects. Essentially, I am comparing two individuals who both spent their adult lives in Canada and who have the same occupation. I find that a majority worker does not necessarily earn more than a minority worker, if both have the average education and experience of the sample. The estimated gap 1.4% is not statistically significant. However, the different payoffs to education and experience remain significant.

Model 3 includes additionally the inverse Mill's ratio to correct for the self-selection of minority workers into the majority labor market. Presumably, this specification will yield more reliable estimates. Notice that the wage gap is reduced to -0.03% and is insignificant. Furthermore, the estimated return to education increases compared with that of Model 2. The difference in payoff to education and experience remains statistically significant.

The reduction of the estimated wage gap after selection correction is puzzling, though the wage gap is not significantly different from zero. We expect that minority workers with higher ability are more likely to enter the

Table 2.4: Within-Labor-Market Wage Gap Ignoring Discrimination

Specification	All CMAs			CMAs outside Quebec		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Majority group D_i	0.1393^a	0.0140	-0.0003	0.1398^a	-0.0025	-0.0109
	(0.0326)	(0.0685)	(0.0747)	(0.0357)	(0.0688)	(0.0739)
Education	0.0629^a	0.0413^a	0.0381^a	0.0600^a	0.0399^a	0.0367^a
	(0.0042)	(0.0033)	(0.0048)	(0.0046)	(0.0036)	(0.0051)
Experience	0.0146^a	0.0179^a	0.0186^a	0.0142^a	0.0177^a	0.0184^a
	(0.0023)	(0.0021)	(0.0022)	(0.0024)	(0.0022)	(0.0023)
Experience sqd.	-0.0001^b	-0.0002^a	-0.0002^a	-0.0001^c	-0.0002^a	-0.0002^a
	(0.0001)	(0.0000)	(0.0000)	(0.0001)	(0.0001)	(0.0001)
(De-meaned Educ)D_i	0.0225^a	0.0115^a	0.0147^a	0.0222^a	0.0110^a	0.0142^a
	(0.0048)	(0.0037)	(0.0051)	(0.0053)	(0.0041)	(0.0052)
(De-meaned Exp)D_i	0.0233^a	0.0143^a	0.0136^a	0.0247^a	0.0157^a	0.0150^a
	(0.0030)	(0.0028)	(0.0029)	(0.0032)	(0.0031)	(0.0030)
(De-meaned Exp^2)D_i	-0.0004^a	-0.0003^a	-0.0003^a	-0.0005^a	-0.0003^a	-0.0003^a
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Sex (Female=1)	-0.1106^a	-0.1136^a	-0.1136^a	-0.1187^a	-0.1166^a	-0.1166^a
	(0.0087)	(0.0077)	(0.0077)	(0.0109)	(0.0095)	(0.0095)
Married	0.1866^a	0.1656^a	0.1657^a	0.1888^a	0.1668^a	0.1670^a
	(0.0102)	(0.0091)	(0.0091)	(0.0119)	(0.0105)	(0.0106)
Sex*Married	-0.0971^a	-0.0955^a	-0.0956^a	-0.0907^a	-0.0889^a	-0.0890^a
	(0.0096)	(0.0106)	(0.0106)	(0.0109)	(0.0115)	(0.0115)
Language score		0.0310	0.0042		0.0460	0.0155
		(0.0606)	(0.0644)		(0.0635)	(0.0666)
Immigrate after 19		-0.1221^a	-0.1183^a		-0.1285^a	-0.1247^a
		(0.0111)	(0.0123)		(0.0134)	(0.0156)
Inverse Mill's ratio			-0.1098			-0.1070
			(0.1268)			(0.1323)
Occ. dummies	No	Yes	Yes	No	Yes	Yes
N	176651	175809	175809	142812	142089	142089
R^2	0.1190	0.1453	0.1453	0.1200	0.1461	0.1461

The dependent variables are log hourly wages. The inverse Mill's ratio is computed using model 2 of Table 2.3 for all CMAs and CMAs outside Quebec, respectively. [a] significance level of 1%. [b] significance level of 5%. [c] significance level of 10%. Standard errors are in parentheses. The regression error terms are clustered by individual language group within a CMA.

majority labor market. Therefore, the uncorrected wage gap is an underestimate of the real wage gap. This presumes a negative correlation between a worker's unobserved ability and her language learning cost.

This result may come from immigration selection. In the data, minority people can be born either in Canada or in another country. The immi-

Table 2.5: Within-Labor-Market Wage Gap Considering Discrimination

Specification	All CMAs			CMAs outside Quebec		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Majority group D_i	0.0687^c	0.0615	0.0476	0.0676	0.0500	0.0425
	(0.0386)	(0.0664)	(0.0692)	(0.0418)	(0.0683)	(0.0709)
Visible minority V_i	-0.1338^a	-0.1146^a	-0.1143^a	-0.1396^a	-0.1157^a	-0.1154^a
	(0.0396)	(0.0244)	(0.0245)	(0.0412)	(0.0265)	(0.0266)
Interaction D_iV_i	0.0280	0.0665^b	0.0648^b	0.0302	0.0651^b	0.0636^c
	(0.0428)	(0.0282)	(0.0285)	(0.0476)	(0.0327)	(0.0331)
Education	0.0624^a	0.0383^a	0.0353^a	0.0596^a	0.0368^a	0.0341^a
	(0.0045)	(0.0043)	(0.0052)	(0.0044)	(0.0041)	(0.0051)
Experience	0.0222^a	0.0208^a	0.0214^a	0.0222^a	0.0210^a	0.0215^a
	(0.0033)	(0.0029)	(0.0031)	(0.0036)	(0.0032)	(0.0034)
Experience sqd.	-0.0002^a	-0.0002^a	-0.0002^a	-0.0002^a	-0.0002^a	-0.0002^a
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
(De-meaned Educ)D_i	0.0230^a	0.0142^a	0.0172^a	0.0224^a	0.0137^a	0.0164^a
	(0.0046)	(0.0040)	(0.0053)	(0.0043)	(0.0037)	(0.0052)
(De-meaned Exp)D_i	0.0164^a	0.0118^a	0.0111^a	0.0177^a	0.0130^a	0.0124^a
	(0.0039)	(0.0034)	(0.0036)	(0.0043)	(0.0038)	(0.0040)
(De-meaned Exp^2)D_i	-0.0003^a	-0.0002^a	-0.0002^a	-0.0004^a	-0.0003^a	-0.0003^a
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
(De-meaned Educ)V_i	0.0033	0.0037	0.0035	0.0034	0.0041	0.0039
	(0.0055)	(0.0050)	(0.0050)	(0.0054)	(0.0050)	(0.0050)
(De-meaned Exp)V_i	-0.0088^b	-0.0048	-0.0048	-0.0092^b	-0.0055	-0.0054
	(0.0040)	(0.0037)	(0.0037)	(0.0044)	(0.0040)	(0.0040)
(De-meaned Exp^2)V_i	0.0001	0.0000	0.0000	0.0001	0.0001	0.0000
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Language score		-0.0617	-0.0873		-0.0507	-0.0773
		(0.0625)	(0.0687)		(0.0673)	(0.0723)
Immigrate after 19		-0.0924^a	-0.0884^a		-0.0968^a	-0.0931^a
		(0.0133)	(0.0139)		(0.0138)	(0.0145)
Inverse Mill's ratio			-0.1063			-0.0950
			(0.1284)			(0.1283)
Occ. Dummies	No	Yes	Yes	No	Yes	Yes
N	176651	175809	175809	142812	142089	142089
R^2	0.1210	0.1459	0.1459	0.1225	0.1468	0.1468

The dependent variables are log hourly wages. Sex, Marital Status, and an interaction of the two are also included as regressors. The inverse Mill's ratio is computed using model 2 of Table 2.3 for all CMAs and CMAs outside Quebec respectively. ^{a} significance level of 1%. ^{b} significance level of 5%. ^{c} significance level of 10%. Standard errors are in parentheses. The regression error terms are clustered by individual language group within a CMA.

grants have undergone an immigration selection process, while the natives have not. Presumably, the immigration process selects people at the upper tail of the ability spectrum using minimum requirement on education, work experience, etc. Therefore, the immigrants have on average higher ability. It is clear natives are more likely to enter the majority market than the immigrants. Therefore, the correlation between unobserved ability and the language learning cost can be positive.

I run specifications in which I exclude those who were born Canadian. Presumably, this will eliminate the concern about immigration selection. The results are shown in Table A.1 in Appendix A.4. The results there are more consistent with the ability sorting story.

Table 2.5 controls for discrimination based on a worker's appearance. I add 'visible minority status' of a worker, its interaction with language group dummy, and its interactions with de-meaned education, experience, and experience squared. The 'visible minority' dummy equals one if a person is a visible minority. I will again only discuss the results for all CMAs since the results for CMAs outside Quebec are similar. Table A.2 shows the same set of specifications for first generation immigrants only.

Model 1, Model 2, and Model 3 have similar interpretations to previous findings. The coefficient before the visible minority dummy is negative, meaning visible minority workers indeed face discrimination. But those who speak the majority language at home face less discrimination, as shown by the positive coefficient before the interaction between D_i and V_i. The message from Table 2.5 is essentially the same as that from Table 2.4. The estimated wage gap is not always significant, but the rewards for education and experience are quite different across the two language groups. I also find that the estimated wage gap is larger for visible minorities, about 10% based on Model 3. The return to experience for visible minorities is also lower.

In summary, the results in Table 2.4 and Table 2.5 are consistent. First, I find a strong difference in returns for education and experience across the language groups. This supports the existence of a Within-Labor-Market Wage Gap. Second, the estimated wage gap for a worker with average characteristics can be positive or close to zero, depending on which conceptual comparison we are making. In my view, Model 3's estimate represents the lower bound and Model 1's represents the upper bound. Finally, visible minorities face higher Within-Labor-Market Wage Gap than non visible minorities.

2.5.3 Within-Language-Group Wage Gap

Table 2.6 presents the evidence for the Within-Language-Group Wage Gap. Recall that it is the wage difference between two comparable minority workers who work in different labor markets. I drop all workers whose home languages are either English or French because of the concern about bilingualism. Again, I present two sets of results for the sample of all CMAs and the sample of CMAs outside Quebec and Ottawa. The second set of results is meant to address the concern about the 'ripple effect'.

I discuss only the results for all CMAs, since the results for CMAs outside Quebec are similar. Model 1 compares one minority worker who was born in Canada and works in the majority labor market versus another minority worker who was born in another country and works in the minority labor market. I find that the one in the majority labor market earns 18.7% more than the one in the minority labor market, both having the average characteristics of the sample. In addition, the return to education and experience is higher for the former. The difference is statistically significant.

Model 2 compares two minority workers who both spent their adult lives in Canada and have the same occupation, but work in different labor markets. I find that a worker in the majority market earns 15.1% more than her minority labor market counterpart, both having the average education and experience of the sample. In addition, the different payoff for education and experience remains statistically significant.

Model 3 includes terms to correct for the self-selection of minority workers into different labor markets. Notice that the wage gap is increased to 36.5%, which is significant at the 10% level. The estimated difference in return to education also increases. The difference in payoff for education and experience remains significant.

The increase in the estimated wage gap looks puzzling. We expect high ability workers to enter the majority labor market and low ability workers to stay in the minority labor market. Therefore, the uncorrected wage gap is an overestimate of the real wage gap. This is not necessarily true. Our discussion in the previous section also applies here. I show in Table A.3 in Appendix A.4 the results excluding those who were born in Canada.

Interestingly, the payoff for education in the minority labor market is quite small, at 1.1%, and not significant according to Model 3. In contrast, the payoff for one year of schooling is 4.3% in the majority labor market. This contrast is reminiscent of the literature on labor market segmentation (Dickens and Lang, 1992), in which workers in different labor market segments receive different returns to education.

Table 2.6: Within-Language-Group Wage Gap Ignoring Discrimination

Specification	All CMAs			CMAs outside Quebec		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Majority market L_i	0.1869^a	0.1512^a	0.3647^c	0.1940^a	0.1541^a	0.3241
	(0.0270)	(0.0250)	(0.2023)	(0.0291)	(0.0274)	(0.2015)
Education	0.0392^a	0.0201^a	0.0111	0.0372^a	0.0190^a	0.0117
	(0.0078)	(0.0058)	(0.0089)	(0.0087)	(0.0063)	(0.0094)
Experience	0.0113^a	0.0163^a	0.0211^a	0.0117^a	0.0164^a	0.0205^a
	(0.0034)	(0.0030)	(0.0038)	(0.0037)	(0.0031)	(0.0039)
Experience Sqd.	-0.0001	-0.0002^a	-0.0003^a	-0.0001	-0.0002^a	-0.0003^a
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
(De-meaned Educ)L_i	0.0253^a	0.0184^a	0.0323^a	0.0244^a	0.0171^a	0.0284^a
	(0.0057)	(0.0051)	(0.0060)	(0.0063)	(0.0056)	(0.0065)
(De-meaned Exp)L_i	0.0073^b	0.0053^b	-0.0008	0.0068^b	0.0049^c	-0.0002
	(0.0028)	(0.0026)	(0.0034)	(0.0029)	(0.0026)	(0.0034)
(De-meaned Exp^2)L_i	-0.0001	-0.0000	0.0001	-0.0001	-0.0000	0.0001
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Sex (Female=1)	-0.1160^a	-0.1099^a	-0.1097^a	-0.1252^a	-0.1107^a	-0.1108^a
	(0.0266)	(0.0253)	(0.0251)	(0.0284)	(0.0267)	(0.0264)
Married	0.0790^a	0.0936^a	0.0925^a	0.0753^a	0.0929^a	0.0918^a
	(0.0226)	(0.0200)	(0.0198)	(0.0240)	(0.0209)	(0.0207)
Sex*Married	-0.0334	-0.0360	-0.0351	-0.0252	-0.0327	-0.0315
	(0.0289)	(0.0234)	(0.0230)	(0.0300)	(0.0238)	(0.0233)
Language Score		0.0860^c	0.0858^c		0.0985^c	0.0991^b
		(0.0494)	(0.0456)		(0.0505)	(0.0430)
Immigrate after 19		-0.1423^a	-0.1495^a		-0.1452^a	-0.1512^a
		(0.0182)	(0.0274)		(0.0195)	(0.0287)
Correction for $L_i = 1$			0.1821			0.1435
			(0.1769)			(0.1731)
Correction for $L_i = 0$			0.2257^b			0.1797
			(0.1134)			(0.1146)
Occ. dummies	No	Yes	Yes	No	Yes	Yes
N	35880	35223	35223	32318	31758	31758
R^2	0.0544	0.0893	0.0899	0.0531	0.0891	0.0896

The dependent variables are log hourly wages. The correction terms are computed using model 2 of Table 2.3 for all CMAs and CMAs outside Quebec respectively. ^{a} significance level of 1%. ^{b} significance level of 5%. ^{c} significance level of 10%. Standard errors are in parentheses. The regression error terms are clustered by individual language group within a CMA.

Table 2.7 shows the result when I control for discrimination against visible minorities. I also allow possible differential discrimination against minority workers across labor markets by adding the interaction of V_i and L_i.

Table 2.7: Within-Language-Group Wage Gap Considering Discrimination

Specification	All CMAs			CMAs outside Quebec		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Majority Market L_i	0.1094^a	0.0848^a	0.0236	0.1041^a	0.0810^a	-0.0538
	(0.0273)	(0.0280)	(0.2106)	(0.0250)	(0.0273)	(0.1910)
Visible Minority V_i	-0.2246^a	-0.2029^a	-0.1634^a	-0.2430^a	-0.2171^a	-0.1910^a
	(0.0468)	(0.0511)	(0.0621)	(0.0455)	(0.0527)	(0.0632)
Interaction L_iV_i	0.0867^b	0.0973^b	0.0526	0.0956^b	0.1041^b	0.0757
	(0.0388)	(0.0421)	(0.0530)	(0.0396)	(0.0440)	(0.0545)
Education	0.0274^a	0.0081	0.0078	0.0236^a	0.0054	0.0076
	(0.0069)	(0.0065)	(0.0083)	(0.0069)	(0.0067)	(0.0084)
Experience	0.0140^a	0.0148^a	0.0169^a	0.0136^a	0.0142^a	0.0154^a
	(0.0045)	(0.0044)	(0.0052)	(0.0046)	(0.0046)	(0.0055)
Experience Sqd.	-0.0001	-0.0001	-0.0002	-0.0001	-0.0001	-0.0001
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
(De-meaned Educ)L_i	0.0253^a	0.0182^a	0.0267^a	0.0256^a	0.0179^a	0.0230^a
	(0.0043)	(0.0042)	(0.0050)	(0.0045)	(0.0046)	(0.0051)
(De-meaned Exp)L_i	0.0076^a	0.0064^b	0.0036	0.0074^b	0.0063^b	0.0048
	(0.0028)	(0.0027)	(0.0037)	(0.0029)	(0.0028)	(0.0039)
(De-meaned Exp^2)L_i	-0.0001	-0.0000	-0.0000	-0.0001	-0.0000	-0.0000
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
(De-meaned Educ)V_i	0.0223^a	0.0203^a	0.0205^a	0.0239^a	0.0218^a	0.0221^a
	(0.0060)	(0.0055)	(0.0055)	(0.0061)	(0.0058)	(0.0058)
(De-meaned Exp)V_i	0.0003	0.0036	0.0028	0.0012	0.0041	0.0033
	(0.0045)	(0.0046)	(0.0047)	(0.0047)	(0.0048)	(0.0049)
(De-meaned Exp^2)V_i	-0.0001	-0.0002	-0.0002	-0.0001	-0.0002	-0.0002
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Correction for $L_i = 1$			0.2723			0.2355
			(0.1794)			(0.1705)
Correction for $L_i = 0$			0.0191			-0.0414
			(0.1125)			(0.1011)
Occ. dummies	No	Yes	Yes	No	Yes	Yes
N	35880	35223	35223	32318	31758	31758
R^2	0.0631	0.0936	0.0937	0.0626	0.0935	0.0937

The dependent variables are log hourly wages. The correction terms are computed using model 2 of Table 2.3 for all CMAs and CMAs outside Quebec respectively. Though not reported in this table, Sex, Marital Status, their interaction, language score, and immigration age dummy are also included as regressors. ^{a} significance level of 1%. ^{b} significance level of 5%. ^{c} significance level of 10%. Standard errors are in parentheses. The regression error terms are clustered by individual language group within a CMA.

Clearly, the results are consistent with those in Table 2.6. The estimated wage gap is mostly positive. The different marginal effect of education is robust.

Based on Model 3, The estimated wage gap for a non visible minority is positive for the sample of all CMAs, at 2.4%, but not significant. The estimated wage gap for a visible minority is higher, at 7.7% = 2.4% + 5.3%. Adding selection corrections reduces the estimated wage gap, which is more in line with the explanation that workers are sorted according to their abilities. This suggests that the change in the wage gap estimate in Table 2.6 may be caused by discrimination. Model 3 also suggests that visible minorities face less discrimination in the majority labor market, represented by the positive coefficient before $V_i \cdot L_i$.[17]

In summary, the results in Table 2.6 and Table 2.7 are clear. First, there exists a Within-Language-Group Wage Gap, at least for visible minorities. Second, there is a robust difference in returns to education and experience across labor markets, suggesting that the urban labor market is indeed segmented.

2.5.4 Comparative Statics

The theoretical model implies a specific relationship between wages and majority population and minority population.[18] I test these relationships in this section.

Table 2.8 reports the wage regressions for minority workers in the majority labor market. Specifically, they speak English at work but speak other languages at home. In Quebec, those are people who speak French at work but speak other languages at home. Again, I exclude workers who speak either English or French at home. I mainly discuss the two 2SLS specifications for all CMAs.

The 2SLS 1 specification handles the endogeneity of the worker's own group's population by two instruments. The first instrument is the distance from the worker's language of origin to her current city of residence. The second instrument is an imputed language group population.[19] The majority population has a positive and significant effect on workers' wages. The minority population however has a negative and insignificant effect. This pattern is consistent with the theory. Specifically, when the population of the majority group doubles, wages increase by 8.9%, which is equivalent to about three additional years of education.

[17] Table A.4 in Appendix A.4 shows the same specifications for first generation immigrants only.

[18] Refer to Proposition 2.3 and the discussion that follows for details.

[19] Please see section 2.4.4 and Appendix A.3 for details.

Table 2.8: Wages of Minority Workers in the Majority Labor Market

Specification	All CMAs			CMAs outside Quebec		
	OLS	2SLS 1	2SLS 2	OLS	2SLS 1	2SLS 2
Log maj. population	-0.0114	0.0897^b	0.0794^c	0.0247	0.0749	0.0799
	(0.0277)	(0.0445)	(0.0446)	(0.0278)	(0.0499)	(0.0502)
Log own group pop.	0.0308^c	-0.0375	-0.0328	0.0101	-0.0235	-0.0258
	(0.0157)	(0.0288)	(0.0288)	(0.0160)	(0.0314)	(0.0316)
Education	0.0275^a	0.0365^a	0.0365^a	0.0218^a	0.0259^a	0.0260^a
	(0.0066)	(0.0082)	(0.0082)	(0.0068)	(0.0077)	(0.0078)
Experience	0.0258^a	0.0258^a	0.0258^a	0.0238^a	0.0239^a	0.0239^a
	(0.0054)	(0.0055)	(0.0055)	(0.0055)	(0.0056)	(0.0056)
Experience Sqd.	-0.0003^b	-0.0003^b	-0.0003^b	-0.0002^b	-0.0002^b	-0.0002^b
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Immigrate after 19	-0.1190^a	-0.1659^a	-0.1650^a	-0.1103^a	-0.1335^a	-0.1344^a
	(0.0265)	(0.0356)	(0.0356)	(0.0249)	(0.0330)	(0.0331)
Inverse Mill's ratio	0.1028	0.4353^c	0.4323^c	0.0087	0.1654	0.1701
	(0.1584)	(0.2284)	(0.2269)	(0.1467)	(0.2039)	(0.2042)
Language dummies	Yes	Yes	Yes	Yes	Yes	Yes
Occupation dummies	Yes	Yes	Yes	Yes	Yes	Yes
N	18801	18801	18801	16948	16948	16948
R^2	0.1014	0.1003	0.1004	0.1000	0.0998	0.0998

The dependent variables are log hourly wages. Though not reported in the table, Sex, Marital Status, Sex × Marital Status, Visible Minority Status(V_i), De-meaned Education × V_i, De-meaned Experience × V_i, and De-meaned Experience Squared × V_i are included as regressors. The inverse Mill's ratio is computed using model 3 of Table 2.3 for all CMAs and CMAs outside Quebec respectively. 2SLS 1 is a specification where only the worker's own language group population is instrumented. 2SLS 2 is a specification where both the majority group population and the worker's own language group population are instrumented. [a] significance level of 1%. [b] significance level of 5%. [c] significance level of 10%. Standard errors are in parentheses. The regression error terms are clustered by individual language group within a CMA.

The 2SLS 2 specification handles the endogeneity of both the worker's own group's population and the majority group population. Two instruments mentioned previously are used for the worker's own group's population. An imputed majority group population is used for the majority group population.[20] The results are essentially the same. The marginal effect of the majority group population on wages is a little lower. I should note that the results for CMAs outside Quebec are not statistically significant, although the signs of the coefficients are correct.

What happens to wages in the majority labor market if the majority population and the minority population both double? This question is closely

[20] Please see section 2.4.4 and Appendix A.3 for details.

Table 2.9: The 1^{st} Stage for Workers in the Majority Market

Specification	All CMAs			CMAs outside Quebec		
	2SLS 1	2SLS 2		2SLS 1	2SLS 2	
Dependent Variable	$\ln(n^m)$	$\ln(n^d)$	$\ln(n^m)$	$\ln(n^m)$	$\ln(n^d)$	$\ln(n^m)$
Log Distance	-2.3587[a]	.0759[a]	-2.2623[a]	-2.1924[a]	.1362[a]	-2.0198[a]
	(.0355)	(.0079)	(.0377)	(.0315)	(.0055)	(.0320)
Predicted $\ln(n^d)$		.9311[a]	1.1290[a]		.9255[a]	1.1797[a]
		(.0013)	(.0060)		(.0009)	(.0053)
Predicted $\ln(n^m)$	.1665[a]	-.0001	.1715[a]	.1468[a]	.0058[a]	.1532[a]
	(.0032)	(.0007)	(.0034)	(.0031)	(.0005)	(.0031)
N	18801	18801	18801	16948	16948	16948
R^2	0.9255	0.9851	0.9162	0.9441	0.9931	0.9421
F Statistic	6299	33600	5547	7723	66247	7431

The dependent variables are log population of the majority group $\ln(n^d)$ or that of the worker's own language group $\ln(n^m)$ depending the setting. 2SLS 1 is a specification where only $\ln(n^m)$ is instrumented. 2SLS 2 is a specification where both $\ln(n^d)$ and $\ln(n^m)$ are instrumented. The distance is measured as that between a worker's current resident city to the representative location the worker's original language group. Predicted $\ln(n^d)$ is the log of the predicted population based on the product of the national majority group population in 2001 and the ratio between the majority group population in a city and the national population of this group in 1991. Predicted $\ln(n^m)$ is calculated similarly. Other independent variables are suppressed for clarity. [a] significance level of 1%. [b] significance level of 5%. [c] significance level of 10%. Standard errors are in parentheses.

related to the Urban Wage Premium, presented by (Glaeser and Mare, 2001). Based on 2SLS 1 estimates, wages increase by 5.2%.

I report the first stage regression in Table 2.9. All the instruments are statistically significant in both 2SLS 1 and 2SLS 2. The R^2's are very high. The F statistics are also high. Clearly, the instruments are relevant in explaining the population variables.

Table 2.10 shows the wage regressions for minority workers in the minority labor market. They speak their home language at work. I again exclude those whose home languages are either English or French.

The 2SLS 1 specification handles the endogeneity of the worker's own group's population. The effect of the majority group population on wages is negative and significant. The effect of the worker's own language group population on wages is positive and significant. These effects are also economically significant. Doubling the worker's own language group's population is equivalent to more than 10 years of experience for a person with 10 years of work experience. Again, the results are consistent with the theory.

The 2SLS 2 specification handles endogeneity of both the worker's own group's population and the majority group population. The estimates are

Table 2.10: Wages of Minority Workers in the Minority Labor Market

Specification	All CMAs			CMAs outside Quebec		
	OLS	2SLS 1	2SLS 2	OLS	2SLS 1	2SLS 2
Log maj. population	-0.0432	-0.1242[a]	-0.1206[a]	-0.0889	-0.1472[b]	-0.1412[b]
	(0.0483)	(0.0436)	(0.0449)	(0.0605)	(0.0577)	(0.0586)
Log own group pop.	0.0802[b]	0.1410[a]	0.1387[a]	0.1034[b]	0.1469[a]	0.1445[a]
	(0.0317)	(0.0335)	(0.0342)	(0.0430)	(0.0455)	(0.0459)
Education	-0.0039	-0.0171	-0.0167	-0.008	-0.0171	-0.017
	(0.0134)	(0.0152)	(0.0153)	(0.0128)	(0.0161)	(0.0160)
Experience	0.0129[b]	0.0134[b]	0.0134[b]	0.0174[b]	0.0175[b]	0.0175[b]
	(0.0058)	(0.0057)	(0.0057)	(0.0069)	(0.0068)	(0.0068)
Experience Sqd.	-0.0001	-0.0001	-0.0001	-0.0002	-0.0002	-0.0002
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Immigrate after 19	0.0283	0.0857	0.0837	0.0431	0.0841	0.0828
	(0.0572)	(0.0589)	(0.0590)	(0.0698)	(0.0766)	(0.0767)
Inverse Mill's ratio	0.4767[c]	0.7524[a]	0.7431[b]	0.4382[c]	0.6211[c]	0.6166[c]
	(0.2420)	(0.2845)	(0.2863)	(0.2608)	(0.3265)	(0.3246)
Language dummies	Yes	Yes	Yes	Yes	Yes	Yes
Occupation dummies	Yes	Yes	Yes	Yes	Yes	Yes
N	7136	7136	7136	6331	6331	6331
R^2	0.0609	0.0601	0.0601	0.0627	0.0624	0.0624

The dependent variables are log hourly wages. Though not reported in the table, Sex, Marital Status, Sex × Marital Status, Visible Minority Status(V_i), De-meaned Education × V_i, De-meaned Experience × V_i, and De-meaned Experience Squared × V_i are included as regressors. The inverse Mill's ratio is computed using model 3 of Table 2.3 for all CMAs and CMAs outside Quebec respectively. 2SLS 1 is a specification where only the worker's own language group population is instrumented. 2SLS 2 is a specification where both the majority group population and the worker's own language group population are instrumented. [a] significance level of 1%. [b] significance level of 5%. [c] significance level of 10%. Standard errors are in parentheses. The regression error terms are clustered by individual language group within a CMA.

similar to those of 2SLS 1. Those estimates suggest that workers in the minority labor market should care more about their own group's population when making their location choices. The payoff for education is extremely low for those workers.

According to the estimates of 2SLS 1, the worker's wages will increase by 1.7% when both populations double. This estimate is much lower than that for workers in the majority labor market. It suggests that workers in the minority labor market do not benefit as much from the positive wage premium that a larger city offers as do workers in the majority labor market. I report the first stage regression in Table 2.11.

In summary, Table 2.8 and Table 2.10 confirm the comparative statics

Table 2.11: The 1^{st} Stage for Workers in the Minority Market

Specification	All CMAs			CMAs outside Quebec		
	2SLS 1	2SLS 2		2SLS 1	2SLS 2	
Dependent Variable	$\ln(n^m)$	$\ln(n^d)$	$\ln(n^m)$	$\ln(n^d)$	$\ln(n^d)$	$\ln(n^m)$
Log Distance	-2.5346[a]	-.1005[a]	-2.5753[a]	-2.0980[a]	.0136[c]	-2.0937[a]
	(.0517)	(.0133)	(.0571)	(.0444)	(.0081)	(.0447)
Predicted $\ln(n^d)$		.9203[a]	.9836[a]		.9181[a]	1.061[a]
		(.0025)	(.0108)		(.0016)	(.0087)
Predicted $\ln(n^m)$	.2442[a]	.0008	.2631[a]	.2156[a]	.0049[a]	.2178[a]
	(.0061)	(.0016)	(.0067)	(.0057)	(.0010)	(.0057)
N	7136	7136	7136	6331	6331	6331
R^2	0.9325	0.9796	0.9191	0.9531	0.9933	0.9525
F Statistic	2651	9204	2180	3459	25348	3411

The dependent variables are log population of the majority group $\ln(n^d)$ or that of the worker's own language group $\ln(n^m)$ depending the setting. 2SLS 1 is a specification where only $\ln(n^m)$ is instrumented. 2SLS 2 is a specification where both $\ln(n^d)$ and $\ln(n^m)$ are instrumented. The distance is measured as that between a worker's current resident city to the representative location the worker's original language group. Predicted $\ln(n^d)$ is the log of the predicted population based on the product of the national majority group population in 2001 and the ratio between the majority group population in a city and the national population of this group in 1991. Predicted $\ln(n^m)$ is calculated similarly. Other independent variables are suppressed for clarity. [a] significance level of 1%. [b] significance level of 5%. [c] significance level of 10%. Standard errors are in parentheses.

of wages implied by the theory. This distinguishes the theory from other competing theories, such as the human capital theory of language skill. Additionally, the results in the tables may explain why new immigrants, who tend to work in ethnic enclaves, cluster in cities where there is a large population of their own language group.

2.6 Robustness and Discussions

2.6.1 Different Wage Measures

I chose log hourly wage as the dependent variable, to be consistent with the labor economics literature. I also measure wages as 'log annual wages' or 'raw annual wages'. These different specifications provide essentially the same results. The tables are available upon request.

Table 2.12: Labor Market Choices of Bilingual Workers

	Outside Quebec		In Quebec	
	Model 1	Model 2	Model 1	Model 2
Education	-0.0513^a	-0.0666^a	-0.0489^a	-0.0506^a
	(0.0143)	(0.0098)	(0.0029)	(0.0023)
Age	-0.0061^a	-0.0064^a	-0.0042^a	-0.0043^a
	(0.0021)	(0.0022)	(0.0001)	(0.0002)
Sex (Female=1)	-0.2549^a	-0.2511^a	0.0170^b	0.0212^a
	(0.0346)	(0.0358)	(0.0067)	(0.0042)
Immigrate after age 19	0.1137	0.0283	-0.1478^a	-0.1169^a
	(0.1320)	(0.0762)	(0.0339)	(0.0098)
Log own group population		-0.3474^a		-0.2954^a
		(0.0527)		(0.0006)
Log majority population		0.3080^a		0.3503^a
		(0.0355)		(0.0009)
Intercept	0.9429^a	0.4825	-0.0529	-1.0017^a
	(0.2975)	(0.6226)	(0.0413)	(0.0298)
N	3871	3871	11755	11755
Pseudo R^2	0.0184	0.0484	0.0106	0.0213

The dependent variable is the worker's language at work. It equals 1 if it is English (or French in Quebec), 0 otherwise. Minority workers are those whose home languages are neither English nor French. ^{a} significance level of 1%. ^{b} significance level of 5%. ^{c} significance level of 10%. Standard errors are in parentheses. The error terms are clustered by individual language group within a CMA.

2.6.2 Bilingualism

Previously, I excluded French speaking workers outside Quebec and English speaking workers in Quebec from the analysis. In this section, I restrict the sample to English speaking or French speaking workers only. I explore whether the bilingual policy in Canada has an effect on those workers' wages. Throughout the chapter, "bilingual" means a person can communicate in both English and French.

Table 2.12 shows the determinants of the labor market choices of English or French speaking workers. The first subsample includes workers who speak French at home and live outside Quebec. They choose their work language: French versus English. The second subsample includes workers who speak English at home and live in Quebec. They also choose their work language: English versus French.

The estimates are not as intuitive as those in Table 2.3. For example, a more educated French-speaking worker is less likely to speak English at work. Outside Quebec, there are a lot of government jobs that require French

Table 2.13: Within-Labor-Market Wage Gap of Bilingual Workers

Specification	Outside Quebec			In Quebec		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Majority group D_i	0.0321	0.0289	-0.0805	0.0037	-0.0037	0.3074^b
	(0.0526)	(0.0478)	(0.1126)	(0.0184)	(0.0156)	(0.0900)
Education	0.0958^a	0.0661^a	0.0715^a	0.0981^a	0.0633^a	0.0538^a
	(0.0104)	(0.0098)	(0.0090)	(0.0059)	(0.0081)	(0.0080)
Experience	0.0317^a	0.0273^a	0.0277^a	0.0254^a	0.0201^a	0.0192^a
	(0.0087)	(0.0085)	(0.0084)	(0.0034)	(0.0035)	(0.0034)
Experience Sqd.	-0.0004^c	-0.0003	-0.0003	-0.0002^b	-0.0002	-0.0002
	(0.0002)	(0.0002)	(0.0002)	(0.0001)	(0.0001)	(0.0001)
(De-meaned Educ)*D_i	-0.0140	-0.0137	-0.0191^b	-0.0099	-0.0098	-0.0003
	(0.0108)	(0.0098)	(0.0090)	(0.0066)	(0.0077)	(0.0077)
(De-meaned Exp)*D_i	0.0065	0.0058	0.0054	0.0065	0.0058	0.0068
	(0.0090)	(0.0087)	(0.0086)	(0.0037)	(0.0037)	(0.0036)
(De-meaned Exp^2)*D_i	-0.0002	-0.0002	-0.0002	-0.0002	-0.0001	-0.0001
	(0.0002)	(0.0002)	(0.0002)	(0.0001)	(0.0001)	(0.0001)
Immigrate after 19		-0.1109^a	-0.1110^a		-0.1361^a	-0.1414^a
		(0.0223)	(0.0223)		(0.0280)	(0.0235)
Inverse Mill's ratio			-0.1312			0.2228^b
			(0.1215)			(0.0596)
Occ. dummies	No	Yes	Yes	No	Yes	Yes
N	119326	118977	118977	26478	26439	26439
R^2	0.1341	0.1578	0.1578	0.1043	0.1321	0.1321

The dependent variables are log hourly wages. In all the specifications, sex, marital status, and the interaction between the two are also included as regressors. ^{a} significance level of 1%. ^{b} significance level of 5%. ^{c} significance level of 10%. Standard errors are in parentheses. The regression error terms are clustered by individual language group within a CMA.

language skills. To the extent that government jobs require higher education, a positive correlation between speaking French at work and education may exist. An English speaking worker in Quebec is also less likely to speak French at work if she is more educated. The effects of populations are nonetheless compatible with the language theory.

Table 2.13 reports the Within-Labor-Market Wage Gap estimates. The two groups to be compared in provinces outside Quebec are English-speaking workers (majority) and French-speaking workers (minority) who speak English at work. The two comparison groups in Quebec are French-speaking workers (majority) and English-speaking workers (minority) who speak French at work. The estimates are mostly insignificant. The differences in the marginal effects of education and experience are also insignificant.

Table 2.14: Within-Language-Group Wage Gap of Bilingual Workers

Specification	Outside Quebec			In Quebec		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Majority market (L_i)	-0.0769^b	-0.0594	0.3475^c	-0.0098	0.0056	-0.3854
	(0.0337)	(0.0347)	(0.1900)	(0.0119)	(0.0123)	(0.1593)
Education	0.0732^a	0.0561^a	0.0635^a	0.0875^a	0.0604^a	0.0573^a
	(0.0099)	(0.0102)	(0.0118)	(0.0042)	(0.0049)	(0.0051)
Experience	0.0402^a	0.0364^a	0.0376^a	0.0371^a	0.0332^a	0.0330^a
	(0.0077)	(0.0072)	(0.0074)	(0.0008)	(0.0006)	(0.0007)
Experience Sqd.	-0.0006^a	-0.0005^a	-0.0005^a	-0.0005^a	-0.0004^a	-0.0004^a
	(0.0002)	(0.0002)	(0.0002)	(0.0000)	(0.0000)	(0.0000)
(De-meaned Educ)*L_i	0.0231	0.0197	0.0247^b	0.0106	0.0083	0.0014
	(0.0136)	(0.0127)	(0.0095)	(0.0102)	(0.0108)	(0.0067)
(De-meaned Exp)*L_i	-0.0070	-0.0064	-0.0070	-0.0120^c	-0.0121^c	-0.0129^c
	(0.0088)	(0.0089)	(0.0093)	(0.0037)	(0.0041)	(0.0043)
(De-meaned Exp^2)*L_i	0.0002	0.0002	0.0002	0.0003	0.0003	0.0003
	(0.0002)	(0.0002)	(0.0002)	(0.0001)	(0.0001)	(0.0001)
Immigrate after 19		-0.2884^b	-0.3069^a		-0.1909^a	-0.2023^a
		(0.1002)	(0.0991)		(0.0087)	(0.0083)
Correction - $L_i = 1$			-0.2833^c			0.2374
			(0.1474)			(0.0997)
Correction - $L_i = 0$			0.2363^b			-0.1568
			(0.1094)			(0.1155)
Occ. dummies	No	Yes	Yes	No	Yes	Yes
N	2959	2938	2938	8912	8824	8824
R^2	0.1212	0.1433	0.1450	0.1082	0.1332	0.1335

The dependent variables are log hourly wages. In all the specifications, sex, marital status, and the interaction between the two are also included as regressors.^{a} significance level of 1%. ^{b} significance level of 5%. ^{c} significance level of 10%. Standard errors are in parentheses. The regression error terms are clustered by individual language group within a CMA.

There are two explanations for the results. First, the language barrier between the English-speaking population and French-speaking population is small enough to have no effect on wages. Secondly, bilingual workers are paid more to compensate for their language skills. This effect therefore neutralizes the wage gap.

Table 2.14 reports the Within-Language-Group Wage Gap estimates. Outside Quebec, I compare two workers who both speak French at home, but one speaks English at work (in the majority market) and the other speaks French at work (in the minority market). In Quebec, I compare two workers who both speak English at home, but one speaks French at work (in the majority market) and the other speaks English at work (in the minority market).

Outside Quebec, Model 1 and Model 2 show negative estimates of the wage gaps, while Model 3 gives a positive estimate. The differences in marginal effects are mostly insignificant. In Quebec, the estimated Within-Language-Group Wage Ggaps are insignificant. The differences in marginal effects are also insignificant.

Table 2.13 and Table 2.14 show that official language speakers are paid similarly no matter which language group they belong to or which work language they choose. This is certainly different from what I found before. The language theory fails to explain this pattern. There is something else affecting those workers' earnings. Is it the bilingual policy? I have to leave the answer to this question to future research.

2.7 Conclusion

This chapter presents a simple model of labor market segmentation in which language plays a significant role. An urban labor market has several segments, in which the languages of business communications are different. The majority labor market has a larger number of workers and firms, so it offers higher matching quality and hence higher wages.

In equilibrium, there are two types of wage gaps. The Within-Labor-Market Wage Gap exists between a majority worker and a minority worker who both work in the majority labor market. The Within-Language-Group Wage Gap exists between two minority workers who work in different labor markets.

Another key implication is the relationship between workers' wages and language group populations. Wages in the majority labor market increase with the majority population due to the market thickness effect. On the other hand, wages in the minority labor market decrease with the majority population, because minority workers leave the minority labor market and enter the majority labor market as the majority labor market becomes thicker.

Wages in the minority labor market increase with the minority population, again due to the market thickness effect. However, wages in the majority labor market may increase or decrease with the minority population. The latter is because the number of minority workers in the majority labor market can either increase or decrease as the minority population increases.

I tested those predictions using 2001 Canadian Census Public Use Micro Data. The reported work language in the data allowed me to identify a worker's labor market. I confirmed the existence of both types of wage gaps. The marginal effects of population measures on wages were consistent with the theory. These results are unique to the language theory of labor market segmentation, and they differentiate the theory from other competing theories.

Chapter 3

Ethnic Diversity and Neighborhood House Prices

3.1 Introduction

I try to answer the following questions in this chapter. First, do people prefer to interact with individuals from similar ethnic background? In other words, does ethnicity affect the formation of personal relationships? Second and more importantly, what spatial pattern do housing prices have if people prefer to live close to those of the same ethnic group?

There are two basic premises underlying the analysis. First, people benefit from social interactions, and the benefits accrued from interactions with people of the same ethnicity are higher than those obtained from interactions with people from a different ethnicity. If intra-ethnicity interactions are preferable to inter-ethnicity interactions, we should see people having friends predominantly belonging to their own ethnic groups.

Second, people have to meet face-to-face to interact effectively, and they incur transportation costs to carry out these meetings. This premise, combined with the first one, implies that people should consider the ethnic composition of a location when deciding where to live. Individuals want to choose a location where they have easy access to all of their friends. Competition for the more advantageous location leads to increase in housing price there. Therefore, the housing price of a location depends on the ethnic group distribution there.

I test the implications of the first premise using the Canadian Ethnic Diversity Survey 2003, which contains detailed information about people's social networks and their ethno-linguistic identities. I find that people's friends are predominantly of the same ethnic group as theirs. This is especially conspicuous for minority people. In addition, language barriers seem to be a major hurdle in forming cross-ethnicity friendships.

I build a model of the housing market based on the above two premises. I provide an example where the house price of a neighborhood is positively correlated with its Herfindahl index of its ethnic groups. In other words, more homogeneous neighborhood commands higher house price.

Many may justifiably argue that people like to live in a more ethnically diverse neighborhood simply to enjoy the various amenities such as unique shops and restaurants. As a result, more ethnically diverse neighborhoods should command higher prices. It is then an empirical matter to see which one of the two opposite effects dominates the other.

I test this implied positive correlation using Census data and housing transactions of the Metropolitan Area of Vancouver in Canada. More specifically, I test whether the house price is positively correlated with the Herfindahl index of a Census Tract.[21] The ethnic diversity literature (see Alesina and LaFerrara, 2005) has long modeled economic performance of a country, a city, or a neighborhood as a function of the Herfindahl index and its variations. Therefore, this chapter is in this tradition, too.

The cross-sectional analysis uses the combined dataset from two sources. The ethnic group composition information is obtained from Canadian Census profile tables for Vancouver in the years of 1986, 1991, 1996, and 2001. The housing transaction data is retrieved from British Columbia Assessment Authority. The advantage of this dataset is its detailed housing characteristics for each transaction.

However, the cross-sectional analysis is problematic in that the Herfindahl index and the house price are simultaneously determined in equilibrium. To address this concern, I construct a panel of census tracts for the years of 1986, 1991, 1996, and 2001 using primarily the Census data. I then exploit the panel structure by running fixed effects regressions and carrying out the conditional difference-in-difference analysis (see Heckman, Ichimura, Smith, and Todd, 1998). The identifying assumptions are less restrictive in both cases than in the cross-sectional regressions. Both methods allow the Herfindahl index to be endogenous to housing price determination to some extent.

All the methods yield essentially the same results. The overall ethnolinguistic Herfindahl index turns out to have mostly insignificant effects on housing prices. Meanwhile, the Herfindahl index of only non-English speaking groups, calculated with the English speaking group being excluded, has a strong positive effect on housing prices. This result suggests that English speaking individuals may prefer to live in a diverse neighborhood, while mi-

[21] Since I use Canadian Census data, all the geographic areas are as defined by Statistics Canada. A typical Census Tract has a population around 4000-5000 and it often includes scores of street blocks.

nority individuals have a strong preference to live close to their own ethnic groups. It is also consistent with the finding that English speaking individuals tend to have a more diverse social network (see Section 3.2).

There is a long list of studies on ethnicity and neighborhood choices. Schelling (1969, 1971) started this line of work. In these models, two ethnic groups, typically black and white, either prefer to live with people of the same group or shun away from people of the opposite group. Models of this sort include Yinger (1976) and Miyao (1978a). They did not substantiate the preference to segregate, which I attempt to do in Section 3.2.

Most segregation papers predict complete segregation, which is hard to match with the reality. Miyao (1978b) analyzed location choices in the presence of ethnic externalities within a discrete choice framework. He showed that taste heterogeneity can lead to stable mixed equilibrium. Bayer, McMillan, and Rueben (2005) extended Miyao (1978b) by incorporating the housing market and housing prices. They estimated the model using microdata on individuals' location choices, personal characteristics, and neighborhood characteristics in the San Francisco Bay Area. However, Bayer, McMillan, and Rueben (2005) are more concerned about individuals' location choices than the determinants of house prices.

In the next section, I analyze whether ethnicity in fact matters in the formation of personal relationships as presumed by the segregation literature. Section 3.3 presents a simple model of social interactions and characterize its effects on housing prices. Section 3.4 describes the housing transaction data and the Census data I use in the empirical analysis. After that, I report the main results on the relationship between housing prices and ethnic group composition of neighborhoods. I then provide additional analysis in Section 3.5 based solely on housing and other socio-economic information from the Census, where I can exploit the panel structure of the dataset. Finally, I conclude and discuss the implications of this chapter in Section 3.6.

3.2 Social Networks and Ethnicity

Interactions within an ethnic group have several advantages. First of all, the benefit from interacting with those of the same ethnic group tends to be higher. This notion is indirectly supported by Alesina and LaFerrara (2002)'s work, where they found that people in racially homogeneous neighborhood are more likely to trust each other. Trust leads to stable long-term relationships, which yield potentially higher benefits. In addition, people from the same ethnic group share similar social and economic constraints, thus they may provide more relevant support. For example, a person referring her friend to a job knows the challenge her friend may face, thus she may provide better assistance than others.

Second of all, people of the same ethnic group often share the same mother tongue. They can communicate more effectively in a language they are most comfortable with. Additionally, there may be tacit language differences across ethnic groups. For example, Sociolinguists find that African American English differs from white English in both its vocabulary and its sentence structures. These differences hinder smooth communications across ethnic groups.

In this section, I focus on the second point. I use a unique data set to test whether language barriers represent a significant factor in affecting people's social networks. The data set is the Canadian Ethnic Diversity Survey (2003) conducted by Statistics Canada and the Department of Canadian Heritage. It contains detailed information on people's social networks and ethnic identities. The target population consists of persons aged 15 and older living in private dwellings in Canada's ten provinces, excluding aboriginal people and people in remote territories.

Figure 3.1 shows the social network composition of English and French speaking people and compares it with that of non-official language speakers. A person reports in the survey how many of her friends or co-workers have the same ethnic identity as hers. Five categories are provided in the survey: all, most, half, a few, and none. Therefore, this question shows how ethnically concentrated a person's social network is.[22]

The ethnic composition of a person's friends, friends until age 15, and coworkers is shown in this figure. Based on individual responses to the question, I calculate the fraction of individuals reporting the five categories respectively. The figure reports the cumulative fraction to make the graph more readable. This describes the fraction of individuals who report to have at least a certain level of concentration, namely the five categories "all", "most", "a half", "a few", and "none". For example, about 15% non-official language speakers report to have all of their friends coming from their own ethnic groups. 45% of them report that at least "most" of their friends share their ethnicity, which means about 30 percent report "most" in the data.

Comparing across ethnic groups, the social networks of minority indi-

[22]In fact, there are two ethnic identities reported in the data: the first ethnicity and the second ethnicity. If a person reports that "most" of her friends have her first ethnicity and "all" of her friends have her second ethnicity, I takes the overall result as "all", which means this person's social network has the highest level of concentration. This logic applies to other cases.

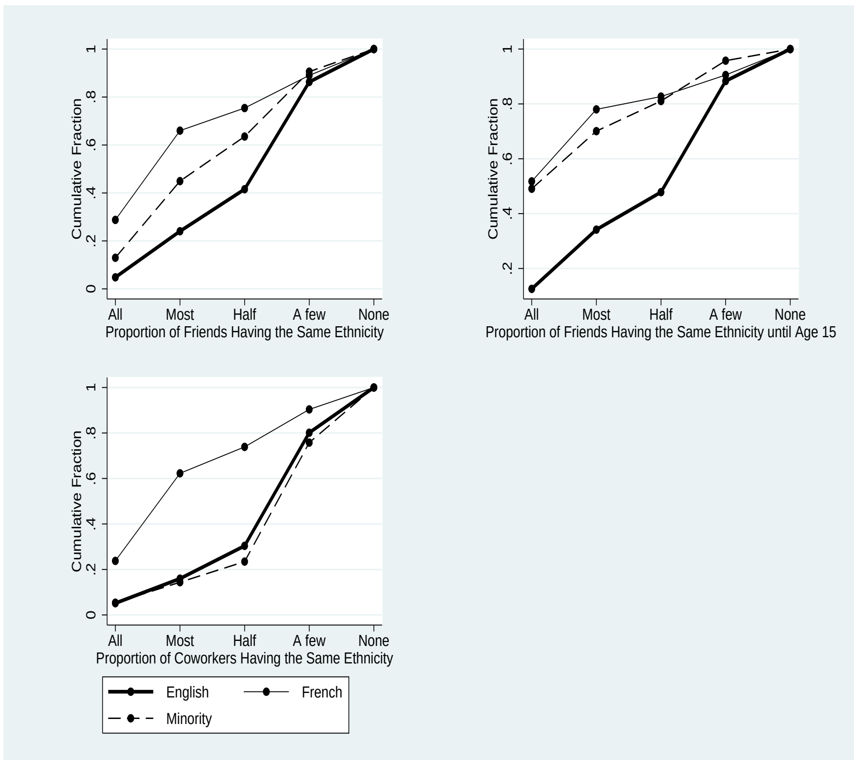

Figure 3.1: Mother Tongue and A Person's Social Network

viduals tend to be more concentrated than those of English speaking individuals. About 64% of them have at least "half" of their friends from the same ethnic group, while the percentage for English speaking people is 42%. Similarly, minority individuals have more ethnically concentrated childhood friends than English speaking people. However, the ethnic composition of minority coworkers is similar to that of English group workers.

Another striking fact shown by this figure is that French speaking individuals tend to have more concentrated social networks. It is even more concentrated than minorities. This may be misleading because the figure does not control for the population share of the ethnic group. In light of this argument, the social network of minorities is in fact much more concentrated than what this figure shows.

To quantitatively determine the relationship between an individual's language skill and her social network, I model the ethnic diversity of an individual's friends as a function of her language skill and a list of control variables. I use individuals' childhood language to proxy for individuals' language skill, which is arguably exogenous to the individual. There are two such proxies: the language used with parents until age 15 and the language used with siblings until age 15. The control variables include education, age, age squared, generation status, the person's language group's population share in the Metropolitan Area, age at immigration, language group fixed effects, and occupation fixed effects.[23]

The results are shown in Table 3.1. The dependent variable is a score of the ethnic diversity of an individual's friends, with "all", "most", "half", "a few", and "none" corresponding to 1-5 respectively. Therefore, it is natural to model this diversity measure as an ordered probit model of a list of explanatory variables. We include only first generation and second generation individuals whose mother tongue is not English or French.[24]

From column 2 to column 4, the table present the estimates when I use childhood language used with parents as a proxy for a person's language skill. In all three specifications, the childhood language with parents has a positive effect on the social network diversity score. Since childhood is important for language learning, those who speak official languages tend to face less language barriers when they become adults. As a result, their friends are more diverse in terms of ethnic composition.

[23]See Appendix B.3 for descriptions of these variables.

[24]Because almost all third generation individuals report either English or French as their mother tongue, there is little variation in terms of their childhood language. However, the effect of childhood language on a person's social network formation is the primary interest of this analysis.

Table 3.1: The Determinants of Soical Networks

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Lan. with Parents	0.6384[a]	0.2316[a]	0.2950[a]			
(Official=1)	(0.0306)	(0.0393)	(0.0558)			
Lan. with Siblings				0.7240[a]	0.3220[a]	0.3430[a]
(Official=1)				(0.0235)	(0.0424)	(0.0614)
Education	0.0442[a]	0.0585[a]	0.0608[a]	0.0444[a]	0.0571[a]	0.0599[a]
	(0.0038)	(0.0048)	(0.0076)	(0.0038)	(0.0048)	(0.0077)
Age	-0.0147[a]	0.0160[a]	0.0372[a]	-0.0043	0.0153[a]	0.0379[a]
	(0.0034)	(0.0044)	(0.0091)	(0.0035)	(0.0045)	(0.0092)
Age^2	0.0002[a]	-0.0001[c]	-0.0003[b]	0.0001[c]	-0.0001[c]	-0.0003[b]
	(0.0000)	(0.0000)	(0.0001)	(0.0000)	(0.0000)	(0.0001)
Generation Status		1.3462[a]	1.2950[a]		1.1666[a]	1.1524[a]
		(0.0773)	(0.1371)		(0.0842)	(0.1460)
Share of Local Pop.		-3.6646[a]	-2.8556[a]		-3.6857[a]	-2.9201[a]
		(0.4045)	(0.5787)		(0.4108)	(0.5877)
Imm. Age Dummies	No	Yes	Yes	No	Yes	Yes
Language Dummies	No	Yes	Yes	No	Yes	Yes
Occ. Dummies	No	No	Yes	No	No	Yes
N	9859	6968	3640	9623	6820	3573
Pseudo R^2	0.0199	0.1104	0.1174	0.0383	0.1128	0.1189

The dependent variable is a score of the ethnic diversity of an individual's friends, with "all", "most", "half", "a few", and "none" of friends sharing the respondents' ethnicity corresponding to 1-5 respectively. The data include only 1st and 2nd generation individuals. [a], [b], and [c] represent significance levels 1%, 5%, and 10% respectively.

The signs of other coefficients are also sensible. Higher education increases the diversity of a person's social relationships. Age has a concave effect on the diversity of a person's social networks. The overall effect tends to be positive. Second generation individuals have a more diverse social networks. The person's ethnic group's share of local population has a negative and significant effect.

The last three columns report the results when childhood language with siblings is used as a proxy for language skills. Childhood language with siblings also has statistically significant effect on the diversity of a person's social networks. The general pattern is similar. Since these language measures are exogenous. The results in this table confirm our presumption that language represents an important barrier for the formation of personal relationships. Therefore, it indirectly supports the notion that intra-ethnicity interactions are preferable to inter-ethnicity interactions.

3.3 Ethnic Preference in the Housing Market

In this section, I build a simple model of the housing market based on two key premises mentioned above. I first present an example to show that the house price and the composition of ethnic groups in a neighborhood can be correlated. After that, I provide a sketch of a neighborhood choice model where house prices and population distribution are endogenized.

3.3.1 Neighborhood House Price and Ethnic Group Composition: An Example

An individual's utility depends on three components: the consumption of the composite good z, the consumption of housing h, and the quality of social interactions q in the neighborhood where she lives. For an individual of ethnic group k, who lives in neighborhood j, her utility function has the following form.

$$U_j^k = (z + q_j^k)^{1-\beta}(h)^\beta, \qquad j = 1, 2, ..., J. \tag{3.1}$$

Note that β is a positive constant between 0 and 1. The budget constraint is $z + r_j h = y^k$, where r_j is rent per unit of housing in the neighborhood. It can be confirmed that the utility is increasing in q_j^k, the quality of social interactions.

I explicitly specify a process whereupon a correlation between the quality of social interactions and the distribution of ethnic groups arises. Each individual is assumed to interact only with those in the same neighborhood. Therefore, she only cares about the local population's ethnic composition. This is of course a great simplification, but the basic results will be similar under more general conditions.

The process is as follows: an individual makes m "visits", in which a person randomly chosen from all the people in the neighborhood acts as her counterpart. We can think of those visits as random encounters in the streets, clubs, and other social settings. The meeting has a quality of b if she meets with a person from the same ethnic group, and a quality of αb otherwise where $0 \leq \alpha < 1$. Therefore, the expected quality of social interactions from living in neighborhood j is

$$q_j^k = mb\Big((1-\alpha)\frac{n_j^k}{N_j} + \alpha\Big), \tag{3.2}$$

where n_j^k is the population of group k in neighborhood j and N_j is the total population in the neighborhood. Clearly, the expected quality of social interactions is increasing in the population share of group k.

I assume that there are only three ethnic groups in the city: a, b, and c. It is straightforward to derive the demand for housing by individuals from these groups respectively. The market clearing condition then implies that the housing price in neighborhood j is

$$r_j = \frac{\beta((y^a + q_j^a)n_j^a + (y^b + q_j^b)n_j^b + (y^c + q_j^c)n_j^c)}{S_j}, \tag{3.3}$$

where S_j is the housing stock in neighborhood j. I assume that S_j is linear in population, or $S_j = \gamma N_j$.

Substituting the expressions for q_j^k and S_j into the above, we get

$$r_j = \frac{\beta}{\gamma}\sum_k[(y^k + mb\alpha)\frac{n_j^k}{N_j}] + mb(1-\alpha)\frac{\beta}{\gamma}\sum_k(\frac{n_j^k}{N_j})^2. \tag{3.4}$$

Note that $\sum_k(\frac{n_j^k}{N_j})^2$ is the Herfindahl index of ethnic groups in neighborhood j.

Proposition 3.1. *If the income is the same across the three groups, then the neighborhood housing price is increasing in the Herfindahl index of ethnic group composition.*

Proof. If $y^a = y^b = y^c \equiv y$, then the first term of the left hand side of Equation (3.4) becomes $(y + mb\alpha)(\beta/\gamma)$. The derivative of r_j with respect to the Herfindahl index $\sum_k (\frac{n_j^k}{N_j})^2$ is then $mb(1-\alpha)(\beta/\gamma) > 0$. Therefore, the neighborhood housing price is increasing in the Herfindahl index. □

Note that social interactions have no effect on housing price if $\alpha = 1$ or if there is no inefficiency for inter-group interactions. Holding others constant, a positive correlation between the Herfindahl index and the neighborhood housing price represents an evidence to support inter-group communication inefficiencies.

If the income is not homogeneous across groups, the relationship between the neighborhood housing price and the Herfindahl index is more complicated. Define the Herfindahl index as $H \equiv \sum_k (\frac{n_j^k}{N_j})^2$.

In addition, we define $R_j^k \equiv n_j^k / N_j$. A natural restriction is then $R_j^a + R_j^b + R_j^c = 1$. Due to this restriction, we have only two degrees of freedom. In other words, we can only vary two of the three populations shares. Equivalently, we can vary one population share and the Herfindahl index. Specifically, we can solve for R_j^b and R_j^c as a function of R_j^a and H.

$$R_j^b, R_j^c = (1/2)[1 - R_j^a \pm (2H - (1 - R_j^a)^2)^{1/2}]. \tag{3.5}$$

Without loss of generalarity, let $R_j^b = (1/2)[1 - R_j^a + (2H - (1 - R_j^a)^2)^{1/2}]$ and $R_j^c = (1/2)[1 - R_j^a - (2H - (1 - R_j^a)^2)^{1/2}]$. This means that group b is larger than group c. After some algebraic manipulations, we get

$$\frac{\partial r_j}{\partial H} = \frac{\beta(y^b - y^c)}{2(2H - (1 - R_j^a)^2)^{1/2}\gamma} + \frac{\beta(1-\alpha)mb}{\gamma}. \tag{3.6}$$

Proposition 3.2. *If the ranking of income across groups and the ranking of group sizes are consistent, the neighborhood house price is increasing the Herfindahl index. Otherwise, as long as the benefit from social interactions b is big enough or the inefficiency from inter-group interactions $1-\alpha$ is big enough, the neighborhood house price is increasing the Herfindahl index.*

Proof. Both claims are directly implied by Equation (3.6). As long as the condition for the first claim is satisfied, both terms in the left hand side of Equation (3.6) are positive. The condition in the second claim can be expressed more formally as $(1-\alpha)mb > \frac{y^c - y^b}{2(2H - (1 - R_j^a)^2)^{1/2}}$. The right hand side of the inequality is always positive, while the left hand side may be negative or positive. □

Throughout the above analysis, we assume that intra-group interactions are always preferable to inter-group interactions. However, we may expect people from the majority group have preference toward living in a diverse neighborhood. For example, Wong (2007) finds that Chinese people in Singapore have inverted U-shaped preference toward living with Chinese. For a neighborhood with higher than 43% Chinese, Chinese in fact prefer to live with people from another ethnic group. In addition, my finding in Section 3.2 also suggests that the English speaking group has a very diverse group of friends, which means they do not necessarily want to live in a homogeneous neighborhood.

In the empirical analysis, I employ two types of Herfindahl index to address the above concern. The first Herfindahl index is measured based on all the ethnic groups, including the English speaking group. The second Herfindahl index is measured based on minority groups only, excluding the English speaking group. I will test whether housing price is positively correlated with these two indices.

3.3.2 A General Model of Neighborhood Choice

In this section, I provide a general model of neighborhood choice based on Miyao (1978b). It is also similar to that presented by Bayer, McMillan, and Rueben (2005).

A number K of ethnic groups live in a closed city. The population of each ethnic group is fixed, denoted by L^k. Each individual chooses one and only one neighborhood in the city. There are a large number of neighborhoods. I denote the number of neighborhoods as J. Each neighborhood has a housing supply of S_j, which is a linear function of population and perfectly divisible. I index neighborhoods by the symbol j.

The indirect utility of an individual who lives in neighborhood j is as follows.

$$\begin{aligned} V^{kj}(r_j, y^k, q_j^k) &\equiv \nu^{kj}(r_j, y^k, q_j^k) + \epsilon_j \\ &\equiv \max_{z,h} u(z, h, q_j^k) + \epsilon_j, \quad \text{s.t. } z + r_j h = y^k. \end{aligned} \tag{3.7}$$

The direct utility function is $u(z, h, q_j^k)$. ν^{kj} is the representative utility of an individual living in neighborhood j. Individual heterogeneity in preference toward neighborhood j is denoted ϵ_j. It can be heterogeneity in preference toward social interactions or toward other neighborhood characteristics.

Since the direct utility and individual heterogeneity is additively separable, the individual decision process can be divided into two steps. First, an individual chooses a consumption plan given a choice of the neighborhood. Second, she compares the maximum utilities obtained from her optimal consumption choices for every neighborhood and chooses the neighborhood offering the highest utility.

To facilitate further analysis, I assume the city is a pure exchange economy. This means the source of individual income comes from individual ownership of neighborhood housing stock and of the composite good. Formally, $y^k = w^k \cdot r$, where $r \equiv (1, r_1, ..., r_J)$ is a vector of prices and $w^k \equiv (w_0^k, w_1^k, ..., w_J^k)$ is a vector of endowments in the composite good and the housing assets in various neighborhoods. Recall that q_j^k is a function of $(n_j^1, ..., n_j^K)$. Therefore, ν^{kj} can be expressed as a function of $(r_1, ..., r_J, n_j^1, ..., n_j^K)$.

Let P^{kj} denote the probability that a group k individual choosing neighborhood j. A randomly drawn individual from the population of group k selects neighborhood j only if it provides the maximum utility. Therefore, P^{kj} can be expressed as

$$P^{kj} \equiv Prob[V^{kj} > V^{ki}] = Prob[\epsilon_j - \epsilon_i < \nu^{ki} - \nu^{kj}] \quad \text{for all } i \neq j. \tag{3.8}$$

If we specify a joint distribution for all $(\epsilon_1, ..., \epsilon_J)$, say joint normal. The selection probability P^{kj} can be expressed as a function of a vector of representative utilities $(\nu^{k1}, ..., \nu^{kJ})$ for group k. Remember that ν^{kj} is a function of r and $(n_j^1, ..., n_j^K)$. Therefore, P^{kj} is a function of r and the vector of population distribution $\mathbf{N} \equiv (n_1^1, ..., n_1^K, ..., n_j^k, ..., n_J^1, ..., n_J^K)$.

Given the distribution of ethnic groups across neighborhoods, the housing market clearing conditions for all the neighborhoods determine the housing prices in all the neighborhoods. In Appendix B.1, I discuss the conditions for existence and uniqueness of such prices. I also characterize these prices in more detail. In the end, the prices are functions of the vector of population distribution $\mathbf{N}$.

The distribution of ethnic groups is then characterized following the idea by Miyao (1978b). The inter-neighborhood equilibrium attains when the number of people from ethnic group k who in fact choose neighborhood j equals the number of people from ethnic group k who are predicted to live in neighborhood j by the discrete choice model above. In Appendix B.2, this problem is reduced to a fixed point problem. I discuss the conditions for existence of the inter-neighborhood equilibrium. Miyao (1978b) provides additional results for uniqueness and stability.

The major difference between this model and that of Miyao (1978b) is the fact that housing prices play a role in people's location choices. The extension establishes that neighborhood housing price is indeed related with the population distribution of ethnic groups. However, this general theory does not specify any specific function form for this dependence. Section 3.3.1 provides such a function form, which will be the subject of our empirical analysis.

3.4 Cross Sectional Analysis

Based on discussions in Section 3.3.1, I can test whether social interactions have material effect on individuals' location choices by regressing housing price on a neighborhood level Herfindahl index of ethnic group concentration. As shown in Section 3.2, minority individuals tend to have friends predominantly from their own ethnic group. To the extent that they value the time spent with friends and they incur transportation costs in interacting with friends, they would choose a location where they have easy access to all of their friends. This leads to a positive correlation between housing price and the Herfindahl index.

The data I use is a combination of detailed housing transactions data and the Census data on socio-economic characteristics of neighborhoods in the Metropolitan Area of Vancouver. The housing transaction data is from British Columbia Assessment Authority (BCAA). I focus primarily on single family homes. It contains information on a house's attributes and its transaction price. The Census data is the 1991, 1996, and 2001 Canadian Census Profile Tables at the Census Tract level.[25]

I combine the two datasets using the street addresses of individual transactions. I match these addresses to the corresponding Census Tracts in Census year 1991, 1996, and 2001. The advantage of combining the two datasets is that both housing and other socio-economic information have good quality.

Table 3.2 reports regressions of log house prices on the Herfindahl Index of a Census Tract's language groups, including the English speaking group. The Herfindahl index measures the overall concentration of ethnic group composition. If there is only one group in the neighborhood, it is equal to 1. If there are more and more language groups, the measure approaches zero. The control variables are of three types: geographic, structural characteristics, and socio-economic characteristics.

Model 1 of Table 3.2 shows the simple correlation between the house price and the Herfindahl index. As a neighborhood becomes more homogeneous, the house price increases on average. The coefficient before distance to

[25] See Appendix B.3 for descriptions of both datasets.

Table 3.2: House Prices and Overall Concentration

	Model 1	Model 2	Model 3	Model 4	Model 5
Distance to CBD	-0.0440[a]	-0.0343[a]	-0.0453[a]	-0.0043[a]	-0.0104[a]
	(0.0004)	(0.0004)	(0.0005)	(0.0004)	(0.0006)
Herfindahl Index	0.1599[a]	0.0482[a]	0.4503[a]	-0.1641[a]	-0.2781[a]
	(0.0128)	(0.0119)	(0.0155)	(0.0141)	(0.0190)
Share of New Imms				0.6031[a]	0.0131
				(0.0193)	(0.0206)
Share of Univ. Grads				1.5376[a]	1.8070[a]
				(0.0419)	(0.0398)
Share of Retirees				1.7118[a]	1.3554[a]
				(0.0384)	(0.0427)
Log Household Inc.				0.5117[a]	0.3354[a]
				(0.0182)	(0.0168)
Share of Apartments				0.0700[a]	-0.0977[a]
				(0.0175)	(0.0175)
Ownership Share				-0.2615[a]	-0.3265[a]
				(0.0270)	(0.0260)
Housing Attributes	No	Yes	Yes	Yes	Yes
Year Dummies	No	No	Yes	No	Yes
Month Dummies	No	No	Yes	No	Yes
Community Dummies	No	No	Yes	No	Yes
N	57,767	57,579	57,579	57,579	57,579
R^2	0.2037	0.3721	0.5372	0.5566	0.6153

The dependent variable is log housing prices. The housing attributes included are the number of bedrooms, number of full bathrooms, number of half bathrooms, age of the house, squared house age, a dummy for the presence of a pool and/or a deck, a dummy for whether it is on a waterfront lot, and dummies indicating the view of the house. The Herfindahl index measures the concentration of all language groups, including the English speaking group. Community fixed effects are dummies for the communities defined by BCAA. These communities are slightly larger than Census Tracts. The standard errors are robust to arbitrary heteroskedasticity. [a], [b], and [c] represent significance levels 1%, 5%, and 10% respectively.

CBD is negative, as expected. As we move away 10 km from the CBD, the house price decrease by 44%. After adding structural controls in Model 2, we observe that the estimates become smaller. The coefficients before the standard structural controls are not reported, but they all have sensible signs.

Model 3 adds year dummies, monthly fixed effects, and community fixed effects. Year dummies are included to address concerns about unobservable time-varying factors that may influence house prices. I include monthly fixed effects to control for possible seasonal variation in housing prices. The community fixed effects are meant to control unobservable community heterogeneity in amenities, school quality, and so on, which may also affect house prices. These communities are defined by BCAA. They are a little larger than Census Tracts. The effect of the Herfindahl index becomes larger.

Model 4 puts in the socio-economic factors, while leaving the year dummies, monthly fixed effects, and community fixed effects aside. The effect of Herfindahl index becomes negative and significant. This suggests that the Herfindahl index are correlated with other socio-economic factors. Because the majority ethnic group tend to have higher average household income and the Herfindahl index may proxy the share of the majority ethnic group, the Herfindahl index and average household income is positively correlated. Therefore, the previous coefficient before the Herfindahl index in Model 1 and Model 2 may be overestimated.

Other variables have sensible signs. Wealthier neighborhood tends to have higher housing values. In addition, retirees who tend to have higher wealth locate in neighborhoods where there are higher house prices. There is also evidence for the existence of human capital externalities: neighborhoods with higher fraction of university graduates tend to have higher house prices. Model 5 adds the time and community fixed effects. The results are basically the same as those from Model 4.

In the following, we analyze the correlation between house prices and the concentration of minority groups, excluding the English speaking population. Table 3.3 presents the cross-sectional regression of log house prices on the Herfindahl index. I trim all observations in Census Tracts whose English speaking population represents more 95% of the total population. This is to exclude the peculiar cases where the measured Herfindahl index is not reliable.

Model 1 of Table 3.3 is again the simplest specification. As the minority groups become more homogeneous the average house price increases. Distance to CBD has the expected sign. Model 2 adds structural character-

Table 3.3: House Prices and Minority Concentration

	Model 1	Model 2	Model 3	Model 4	Model 5
Distance to CBD	-0.0353^a	-0.0291^a	-0.0354^a	-0.0062^a	-0.0137^a
	(0.0004)	(0.0004)	(0.0005)	(0.0004)	(0.0005)
Herfindahl Index	0.4572^a	0.3917^a	0.1682^a	0.0783^a	0.2519^a
	(0.0133)	(0.0129)	(0.0135)	(0.0119)	(0.0138)
Share of New Imms				0.6538^a	0.0028
				(0.0168)	(0.0193)
Share of Univ. Grads				1.4442^a	1.6679^a
				(0.0396)	(0.0381)
Share of Retirees				1.6969^a	1.2447^a
				(0.0377)	(0.0431)
Log Household Inc.				0.5187^a	0.3277^a
				(0.0181)	(0.0169)
Share of Apartments				0.0770^a	-0.0784^a
				(0.0178)	(0.0177)
Ownership Share				-0.2828^a	-0.3551^a
				(0.0270)	(0.0260)
Housing Attributes	No	Yes	Yes	Yes	Yes
Year Dummies	No	No	Yes	No	Yes
Month Dummies	No	No	Yes	No	Yes
Community Dummies	No	No	Yes	No	Yes
N	57,767	57,579	57,579	57,579	57,579
R^2	0.2187	0.3833	0.5305	0.5559	0.6159

The dependent variable is log housing prices. The housing attributes included are the number of bedrooms, number of full bathrooms, number of half bathrooms, age of the house, squared house age, a dummy for the presence of a pool and/or a deck, a dummy for whether it is on a waterfront lot, and dummies indicating the view of the house. The Herfindahl index measures the concentration of all minority language groups, excluding the English speaking group. Community fixed effects are dummies for the communities defined by BCAA. These communities are slightly larger than Census Tracts. The standard errors are robust to arbitrary heteroskedasticity. ^{a}, ^{b}, and ^{c} represent significance levels 1%, 5%, and 10% respectively.

istics. The Herfindahl index coefficient becomes smaller. After I include the time and community fixed effects in Model 3, the coefficient becomes even smaller, but it is still significant. Model 4 includes the socio-economic variables, while leaving the time and community fixed effects out. The effect of the Herfindahl index becomes much smaller, yet still significant at 1% level.

Model 5 is the preferred specification. It controls both the structural characteristics and the socio-economic characteristics. As the minority groups become more concentrated toward one particular group, the house price increases. All the other coefficients are intuitively appealing.

In summary, the combined dataset provides strong evidence for a positive correlation between minority group homogeneity and house prices. The overall concentration of language groups seems to be positively correlated with house prices. However, this correlation reverses its sign once we control some socio-economic variables. The overall Herfindahl index has at best an ambiguous effects.

However, we should interpret those results with caution. First, the theory shows that the Herfindahl index and the house price are co-determined in equilibrium. Therefore, the Herfindahl index may be endogenous. In addition, there is always the possibility of an omitted variable problem. The omitted variable may be correlated with not only the Herfindahl index but also other included variables. This may lead to inconsistent estimates. This problem is somewhat mitigated because I have good housing attributes data and I include both time and community fixed effects. However, it is still a serious concern. In the next section, I employ panel data techniques to alleviate both concerns.

3.5 Panel Data Analysis

3.5.1 Fixed Effects Regressions

I construct a panel of Census Tracts in Vancouver Census Metropolitan Area for Census year 1986, 1991, 1996, and 2001. The data source is Canadian Census profile tables for the corresponding Census years.

The advantage of panel data is that we can employ the panel structure to account for an unobservable census tract specific house price heterogeneity. It is allowed to be arbitrarily correlated with the regressors, including the Herfindahl index. In addition, these unobservable fixed effects can vary arbitrarily across census tracts. This allows an arbitrary house price surface within a city, which is superior to modeling house price by the distance to a presumed CBD.

Table 3.4: House Prices and Overall Concentration - Fixed Effects

	Model 1	Model 2	Model 3	Model 4
Herfindahl Index	.0386	-.0980	-.0647	-.1130^c
	(.0788)	(.0654)	(.0825)	(.0641)
House Age		.0013		.0013
		(.0015)		(.0014)
No. of Rooms		.1805^a		.1631^a
		(.0169)		(.0168)
Share of Apartment		-.0631		-.0751^b
		(.0396)		(.0368)
Ownership Share		-.9651^a		-1.0114^a
		(.0656)		(.0646)
Share of New Immigrants			-.0276	.0194
			(.0688)	(.0555)
Share of Univ. Grads			-.2982^c	.2922^c
			(.1662)	(.1595)
Share of Retirees			.4896^b	.6769^a
			(.2409)	(.1791)
Log Household Inc.			.2532^a	.1267^c
			(.0821)	(.0749)
Year Dummies	Yes	Yes	Yes	Yes
N	1,543	1,543	1,541	1,541
R^2	0.5182	0.6552	0.6068	0.7530

The dependent variable is log average housing value. Each observation represents one census tract in Vancouver Census Metropolitan Area. The Herfindahl index is measured in terms of all language groups, including the English speaking group. ^{a}, ^{b}, and ^{c} represent significance levels 1%, 5%, and 10% respectively.

The disadvantage of this approach is that I can no longer use real transaction price as I did in the last section. First, Census only reports the the owner-reported house prices. Second, the transaction price data is not rich enough to construct a time series of average housing prices at the Census Tract level. Another disadvantage is that the structural controls are less precise in Census.

Table 3.4 reports the fixed effects regression results in a model where log house price is a linear function of a neighborhood's ethnic group Herfindahl index and other attributes.[26] In the first three models, the ethnic group Herfindahl index turns out to be insignificant. The signs of the control variables are broadly consistent with those in Table 3.2. Our baseline model Model 4 finds a negative effect of ethnic group concentration index. This

[26]Note that I drop the the time-invariant distance to CBD variable.

Table 3.5: House Prices and Minority Concentration - Fixed Effects

	Model 1	Model 2	Model 3	Model 4
Herfindahl Index	.1315^b	.1492^a	.2898^a	.2246^a
	(.0587)	(.0480)	(.0481)	(.0373)
House Age		.0019		.0021
		(.0015)		(.0014)
No. of Rooms		.1831^a		.1605^a
		(.0163)		(.0164)
Share of Apartment		-.0553		-.0681^b
		(.0387)		(.0359)
Ownership Share		-.9418^a		-.9777^a
		(.0658)		(.0649)
Share of New Immigrants			-.1065^c	-.0108
			(.0615)	(.0507)
Share of Univ. Grads			-.3393^b	.2607^c
			(.1645)	(.1562)
Share of Retirees			.4727^b	.6625^a
			(.2363)	(.1753)
Log Household Inc.			.2945^a	.1599^b
			(.0851)	(.0771)
Year Dummies	Yes	Yes	Yes	Yes
N	1,543	1,543	1,541	1,541
R^2	0.5337	0.6739	0.6392	0.7828

The dependent variable is log average housing value. Each observation represents one census tract in Vancouver Census Metropolitan Area. The Hefindahl index is measured in terms of minority language groups only, excluding the English speaking group. ^{a}, ^{b}, and ^{c} represent significance levels 1%, 5%, and 10% respectively.

result is consistent with that presented in Model 5 of Table 3.2, though the estimated coefficient is much lower.

Table 3.5 presents the fixed effect regressions that relate house prices to a neighborhood's minority ethnic group concentration, excluding the majority group. In all the specifications, the Herfindahl index for minority groups has a positive effect on house prices. In Model 4, as a neighborhood changes from infinitely many minority groups to only one group (the Hefindahl index changes from 0 to 1), the average house price increases by 22%. This estimate is a little lower than the previous estimate of 25% in Table 3.3 Model 5. The estimated coefficients for socio-economic factors change a lot from Model 5 of Table 3.3, suggesting that the unobservable fixed effect may be correlated with those variables.

In this section, I employ panel data methods to test the relationship between ethnic group homogeneity and house prices. I find a positive relationship between minority group homogeneity and house prices. On the other hand, there seems to be an ambiguous relationship between overall ethnic group homogeneity and house prices. Those results confirm the previous findings using the cross-sectional regressions which rely on far more restrictive identifying assumptions.

The panel data method partially resolves the endogeneity issue in cross-sectional regressions. It does have its own limitations. First, the unobservable fixed effect must be time-invariant. Therefore, some time-variant factors may still be correlated with the regressors. Second, the linear specification is restrictive. More specifically, it restricts the house price to be a linear function of the Herfindahl index. In the following, we adopt a different approach - the conditional difference-in-difference approach, which relies on less restrictive assumptions.

3.5.2 Conditional Difference-in-Difference Analysis

The conditional difference-in-difference method is first proposed by Heckman, Ichimura, Smith, and Todd (1998). This method combines propensity score matching with difference-in-difference method. The propensity score matching method is first proposed by Rosenbaum and Rubin (1983) to nonparametrically identify the treatment effect. The difference-in-difference method has long been used in the natural experiment literature and other program evaluation papers.

The basic idea of propensity score matching is to first model the treatment as a Probit or Logit model. After obtaining a predicted probability of treatment, the so-called "propensity score", the researcher identifies the matching observation for a treatment observation based on the distance between the pair in terms of the "score". The researcher can choose the closest neighbor, the average of several close neighbors, or adopt an arbitrary range of distance to find the matched observations. The identifying assumption is that the treatment status is independent of the potential outcome once the propensity score is conditional upon, the so-called Conditional Independence Assumption.

The difference-in-difference method is similar to fixed effect regressions. What is special is that the variable of interest is a dummy treatment variable. It basically assumes that the treated observation has a common trend as the untreated observation so that the difference in time changes of the outcome between those two observation should be the treatment effect. Including additional controls is equivalent to adding regressors in a fixed effect regression. Then the assumption becomes less restrictive in that it is now conditional upon the observed control variables. However, it is based on restrictive parametric assumptions, namely that the fixed effect regression is a correct specification of the actual outcome.

The advantages of the conditional difference-in-difference methods are two-fold. First, it inherits the benefit of propensity score matching which does not rely on particular function form assumptions about the outcome equation. As a result, the method is superior to fixed effect regressions or traditional difference-in-difference method. Second, it relaxes the Conditional Independence Assumption by allowing possible dependence between treatment status and conditional outcome. Instead, it assumes that the bias, the difference between the counterfactual of the treatment group and the matching nontreatment observed outcome, is constant over time. Therefore, we can difference out the bias by using time differences.

To exploit this conditional difference-in-difference method, I have to discretize the changes in the ethnic group concentration variable. I rank all the census tracts in Vancouver Census Metropolitan Area into 10 deciles according to their ethnic group concentration indices. Depending on the observation date, a census tract may fall in different deciles. The treatment is then defined as moving up at least one rank along the 10 deciles, while all other cases are defined as nontreatment.[27]

The dataset is again a panel constructed for Vancouver Census Metropolitan Area for census year 1986, 1991, 1996, and 2001. We have multiple before-after observations rather than the usual one set of before-after observations in traditional difference-in-difference implementation. I assume that each before-after pair can be treated as independent.[28] Therefore, I can treat all the observations as if they do not have time indices.

I use Probit to model the treatment. However, I do not report the estimated Probit models here. Table 3.6 reports the estimated average treatment effect on the treated for various cases. I present the estimates using all the observed before-after pairs and each time interval 1986-1991, 1991-1996, and 1996-2001 separately. Two ways of matching are used. The first approach selects the census tract that has the closest propensity score (or the predicted probability for treatment) to the treated census tract. The second approach instead chooses the 10 closest census tracts in terms of the propensity scores, averages all the outcomes, and then computes the difference-in-difference estimate.

The first panel of Table 3.6 shows the relationship between overall ethnic group composition and the housing value changes. The interpretation of

[27] I also try dividing the census tracts to 5 and 20 quantiles. The results are similar.

[28] The assumptions made here are formally presented in Miquel (2003).

Table 3.6: Conditional Difference-in-Difference Estimates

		All Years	1986-1991	1991-1996	1996-2001
Herfindahl	Best Match	-.0236^b	-.0243	-.0480^a	-.00329
Index 1		(.0111)	(.0206)	(.0182)	(.0187)
	10 Best Matches	-.0220^b	-.0351^b	-.0363^c	.0009
		(.0096)	(.0170)	(.0188)	(.0146)
	Number of Obs.	314	101	95	118
Herfindahl	Best Match	.0364^a	.0406^a	.0588^a	.0101
Index 2		(.0102)	(.0147)	(.0211)	(.0167)
	10 Best Matches	.0283^a	.0357^a	.0275^b	.0217
		(.0074659)	(.0125)	(.0124)	(.0139)
	Number of Obs.	369	126	120	123

The reported values are calculated treatment effect of ethnic group concentration increase on log housing values. Herfindahl Index 1 means that treatment is defined as a ranking increase in terms of a Census tract's Herfindahl Indices of all language groups. Herfindahl Index 2 means that treatment is defined as a ranking increase in terms of a Census tract's Herfindahl Indices of its minority language groups. Two methods of matching are employed: best match and the average of 10 best matches. ^{a}, ^{b}, and ^{c} represent significance levels 1%, 5%, and 10% respectively.

those estimates is as follows. Based on the "best match" method, the average decrease in housing value is 2.36% for those census tracts that become relatively more homogeneous compared with the potential housing value changes had they became relatively less homogeneous. Both the "best match" approach and the "10 best matches" approach provide negative treatment effect estimates. In addition, the results for separate time intervals are not always in the same direction or statistically significant. For example, the estimate using the "best match" method during the interval 1986-1991 is -2.43% and not significant.

The second panel of Table 3.6 presents the same set of results relating minority ethnic group composition to the house value changes. Similar interpretation applies here, too. Based on "best match" method, the average treatment effect on the treated is a 3.64% increase in housing values over five years. The same estimate using "10 best matches" method is a bit lower at 2.83%. Different from the results in the first panel, the results of separate years are generally consistent with the overall results. The exception is the time interval 1996-2001, where the estimates are not statistically significant though they are still positive.

Generally speaking, those results from the conditional difference-in-difference method are consistent with previous cross-sectional and panel data results. As the minority ethnic group composition becomes more homogeneous, the house prices of the neighborhood increases. It should be noted, however, that the interpretation of those results are not the same as before. The difference is due to the discretization of the ethnic group index into a binary treatment variable.

Table 3.7 adopts a variation of the previous conditional difference-in-difference method. I focus on the neighboring tracts of a particular treatment tract. From the neighboring tracts, I choose the nontreatment tract that has the closest propensity score to that of the treatment tract as the matching tract. This method is called "neighbor and best match" in the table. I also obtain an average change of all nontreatment tracts that are also neighbors of the treatment tract. After that, I compute the difference-in-difference estimate as before. This method is called "average of all neighbors" in the table.

The rationale for the above methods is as follows. Though we can carry out the matching task by focusing on the observable characteristics, the validity of this method hinges on the assumption that the bias conditional on all the observables is constant over time. If this condition is violated, e.g., there is not enough conditioning variables in the conditioning set, the previous result may be unreliable.

One way to resolve this concern is to focus on the neighboring tracts. Suppose two neighboring tracts have in common some unobservable characteristics, which in turn affect whether a tract is subject to treatment or not not, selecting among the neighboring tracts essentially conditions on these unobservable characteristics. This slight revision of conditional difference-in-difference methodology provides a different set of results that is robust to the previous concern. However, the cost is that we only have limited candidates to select from. The estimates may be more noisy as a result.

The results shown in the first panel of Table 3.7 are about the relationship between overall language group composition and house price changes. According to the "neighbor and best match" method, the overall average treatment effect is -1.17% and it is not statistically significant. The other method "average of all neighbors" calculates an overall estimate of -1.92%, which is statistically significant at 10% level. However, the results based on separate intervals are not always in the same direction or statistically significant. Therefore, the results about the effect of overall ethnic group composition on house prices are mixed.

The second panel of Table 3.7 shows the relationship between minority ethnic group composition and house price changes. The results are much stronger than those in the previous panel. Both methods produce similar estimates based on all the time intervals and specific time intervals. The treatment effect of a census tract becoming relatively more homogeneous in

Table 3.7: Spatial Conditional Difference-in-Difference Estimates

		All Years	1986-1991	1991-1996	1996-2001
Herfindahl Index 1	Neighbor + Best Match	-.0117	-.0474[b]	.0120	-.0048
		(.0117)	(.0207)	(.0191)	(.0199)
	Average of Neighbors	-.0192[c]	-.0656[a]	.0148	-.0126
		(.0099)	(-.0656)	(.0152)	(.0153)
	Number of Obs.	286	81	88	117
Herfindahl Index 2	Neighbor + Best Match	.0264[a]	.0372[a]	.0154	.0271
		(.0095)	(.0127)	(.0180)	(.0176)
	Average of Neighbors	.0165[b]	.0217[b]	.0002	.0274[c]
		(.0081)	(.0105)	(.0153)	(.0151)
	Number of Obs.	344	109	114	121

The reported values are calculated treatment effect of ethnic group concentration increase on log housing values. Herfindahl Index 1 means that treatment is defined as a ranking increase in terms of a Census tract's Herfindahl Indices of all its language groups. Herfindahl Index 2 means that treatment is defined as a ranking increase in terms of a Census tract's Herfindahl Indices of its minority language groups. Two methods of matching are employed: neighbor and best match and the average of all nontreated neighbors. [a], [b], and [c] represent significance levels 1%, 5%, and 10% respectively.

terms of its minority ethnic group composition has a positive effect about 2.64% according to "neighbor and best match" method and 1.65% according to "average of all neighbors" method.

To summarize, I find that neighborhoods that experience relative increase in its minority ethnic group concentration also experience higher house price appreciation. In fact, we estimate an average treatment effect of this particular kind of "treatment", so we've inferred from the data the causal effect of minority ethnic group concentration on house prices. However, the effect of overall ethnic group concentration has an ambiguous effect on house prices.

This answers our primary question about how to estimate people's valuation on social interactions. To the extent that minority people tend to have friends from the same ethnic group, they would put more value to a particular location where they can find better social interaction opportunities. On the average, more homogeneous neighborhoods offer better social interactions quality.

3.6 Conclusion

I build a model of neighborhood choice in which housing prices are endogenously determined. People prefer to live with those of the same ethnic background primarily for social interactions benefits. The equilibrium house price is related to the composition of ethnic groups in a neighborhood because people are willing to pay more for neighborhoods where they can have higher quality social interactions. Under some additional conditions, the model implies that house prices increase if the neighborhood becomes more homogeneous in terms of ethnic group composition.

To explore whether people do prefer to interact within ethnic boundary, I analyze people's social networks as they are related with their ethnic belongings using Canadian Ethnic Diversity Survey (2002). I find that minority people tend to have friends predominantly from their own ethnic group. Language barriers are important factors in affecting the ethnic composition of a person's social networks.

I follow the ethnic diversity literature by exploring a direct relationship between the house price and the concentration of ethnic groups, measured by the Herfindahl index. I employ cross-sectional regression, fixed effect regression, and a nonparametric conditional difference-in-difference method to test this relationship.

These three methods have their own merits and limitations. The cross-sectional regression uses a dataset whose quality is the best. By geocoding individual transactions, I am able to combine the housing transaction data with Census data. The dataset therefore has both detailed and accurate information of housing characteristics and socio-economic variables of the neighborhoods. However, it may suffer from omitted variable problem.

The panel data method partially mitigates the omitted variable problem. However, its restrictive parametric form limits its ability to infer a causal effect of ethnic group composition on housing values. The conditional difference-in-difference method does not rely on particular parametric form of the housing price equation. It also has the advantage of intuitive interpretations.

All three methods yield essentially the same results. There is a positive relationship between minority ethnic group concentration and housing prices. On the other hand, the effect of overall ethnic group concentration has at best ambiguous effects on neighborhood housing prices.

Chapter 4

Informed Speculation about Trading Flows

4.1 Introduction

The role of speculators in price determination has attracted a lot of attention by economists. Many believe that speculators stabilize prices. The classical argument asserts that speculators buy low and sell high, and hence smooth the price movement. Counter examples have been offered by Baumol (1957), Telser (1959), and Farrell (1966) under assumptions that speculators are non-competitive or non-speculators have irrational expectations. Hart and Kreps (1986) relax these assumptions and still find that speculation is price-destabilizing.

This chapter analyzes speculation and price formation in the housing market. More specifically, I explore properties of housing price when some traders possess superior knowledge of future price. I define those traders as informed traders. Other traders who do not have such knowledge can however infer from the price of the informed traders' private information. I call those uninformed traders.

In the model, price varies over time because of the random housing demands by noise traders. Noise traders are those whose demands for housing are exogenous. One example would be those who recently move to a city. They have to find a place to live, and there is a minimum level of space needed for each of them. Another example would be those who leave the city. They have to sell their houses, the sizes of which are determined some time in the past. Their housing demands, positive or negative, have some exogenous elements.

First, migrants change the demographics of the city. Many housing economists believe that demographics determine household formation, and the number of households in turn determines the aggregate housing demand. For example, Mankiw and Weil (1989) try to relate housing price trends to demographic changes of US population in 1970's and 1990's.

Second, migrants change the population of a city. There are both theoretical and empirical support for the positive relation between population growth and housing prices. Capozza and Helsley (1989) present a model in which expected population growth is capitalized into land price as well as housing price. Moreover, Capozza, Hendershott, and Mack (2004) find empirically significant effect of population growth on equilibrium housing price.

Last, migrants add noise to the prices. The effect of net migration on the demographic composition of a city is complicated and random, so it may induce random variation in housing demand and hence price variation. Moreover, since the net migration to a city is itself random, its effect on population growth and hence equilibrium price is also uncertain. Therefore, these migrants can be thought of as noise traders.

The supply side of the housing market is mute in the above discussion. I can incorporate housing supply into the model by redefining the demands from noise traders as the demands by migrants net of new housing supply. Here, I implicitly assume that supply is independent of price just as noise traders' demands are. This assumption is restrictive, but it allows me to focus on speculation instead of being distracted to modeling housing supply.

In the model, some people are informed about the net housing demand by noise traders, while some are not. Since we know the various components of noise traders' net housing demand, we get a better understanding about what the private information may be. It can be that some traders are better able to analyze the demographic trend of a city. It can be that some people understand better how demographics affect individuals' housing choices. It is also possible some investors have better knowledge about the future supply of housing in a city or a neighborhood.

The model extends the model by Grossman and Stiglitz (1980). There are four periods. At the beginning, each trader is endowed with both housing asset and riskless bond. At $t = 1$, informed traders receive signals about noise traders' housing demand at $t = 2$ and take actions. Uninformed also change their positions because they observe the price and expect the price to reflect informed traders' new signals. At $t = 2$, all the traders trade again. At $t = 3$, the uncertain payoff of housing asset is paid to each trader, who consumes her wealth. I should note that the signals at $t = 1$ is assumed to be independent of the payoff at $t = 3$, so the signals are purely about random price variation.

This model is able to explain the two puzzling phenomena of housing price series: serially correlated return and return volatility clustering. Serial correlation between consecutive returns happens because the equilibrium

Figure 4.1: Time Line of the Model

price cannot incorporate the new information instantly since not all traders receive the new information. Return becomes more volatile in the period when speculation happens and it becomes less volatile in the period immediately after speculation. This result shows that speculation can both stabilize and destabilize the price, depending on which period we are interested in.

The model contains several comparative statics for autocorrelation and volatility with respect to model parameters: investor risk aversion, the precision of the signal, and the fraction of informed investors. I also calculate the transaction volume in each date to explore the effect of speculation on market transactions. Comparative statics of transaction volume with respect to all the model parameters are also presented. Moreover, an extension of the basic model by imposing non-negativity constraints on asset holdings can explain the puzzling price-volume correlation(see for example Stein, 1995) in the housing market.

The rest of the chapter contains three sections. First, I introduce the basic setup of the rational expectations model. Next, I derive the equilibrium price at each trading date and explore basic properties of the equilibrium using a numerical example. After that, three separate comparative statics are explored in more detail, namely those of return autocorrelation, volatility, and transaction volume. Last, I conclude and summarize the results of the model.

4.2 Model Setup

4.2.1 Time Line of the Model

A pure exchange economy, a country or a city, consists of three groups of traders - informed traders, uninformed traders, and noise traders, who trade on a single risky asset - housing. Both informed traders and uninformed traders belong to the group of rational traders. Rational traders decide their trading strategy given their preferences and information, while noise traders' demand is independent of the price and of the available information. The traders can also invest in a riskless bond which pays zero interest rate.

Figure 4.1 shows the time line of the model. At the beginning of $t = 0$, each trader is endowed with a fixed amount of riskless bond $\bar{b}$ and the risky asset $\bar{z}$. Trading of the single risky asset occurs at $t = 1$ and $t = 2$. The risky asset's dividend is paid off at $t = 3$, after which each trader consumes his/her terminal wealth. Borrowing or lending is only constrained by traders' initial endowment.

Throughout the time horizon, all traders have common beliefs about the final dividend payment of one unit of the risky asset at $t = 3$. This random dividend $\tilde{d}$ is Normal with mean μ and variance σ^2. The exogenous but random trading positions of noise traders at $t = 1$ and $t = 2$ induce random price variations in those periods. In both periods, there are zero net supplies of the risky asset. Price adjusts at $t = 1$ and $t = 2$ to clear market in each period.

Initially at $t = 0$, it is common knowledge that noise traders' random demand at $t = 1$ and $t = 2$ are normal random variables. To facilitate future derivation, I normalize noise traders' demand by the total number of rational traders. These normalized random demands are denoted $\tilde{z}_{n1}$ and $\tilde{z}_{n2}$, each being Normal with mean 0 and variance σ_z^2.

At the beginning of period 1, a fraction λ of rational traders receive the same signal about noise traders' period 2 demand $\tilde{z}_{n2}$. Those traders are informed traders. I assume $\tilde{z}_{n2} = y + \epsilon$, where $y \sim N(0, \sigma_y^2)$ and $\epsilon \sim N(0, \sigma_\epsilon^2)$ are independent random components of $\tilde{z}_{n2}$. The private signal enables the informed traders to observe y, but not to observe ϵ. Note that $\sigma_y^2 + \sigma_\epsilon^2 = \sigma_z^2$

The rational traders that do not receive such signals are classified as uninformed traders. Uninformed traders know the fraction of informed traders and the fact that informed traders have received private signals. They can partially infer from the price the informed traders' private information. This inference is not perfect because the price is noisy due to the trading by noise traders.

In period 2, all the participants submit their demands, resulting in a price that clears the market. At $t = 3$, each trader is paid according to their holdings of the risky asset and the riskless bond. They then consume their terminal wealth.

In the context of housing market, we can think of the private signal as information about future migration flows to a city. The migrants or the noise traders enter the market only at $t = 2$. The traders who live in the city at $t = 1$ expect the migrants to come, but they have different beliefs about the uncertain flow. Some observe the signal, while others do not. They both speculate on the future prospect of housing price and adjust their holdings. At $t = 2$, new migrants come and the speculators sell their holdings to the newcomers.

4.2.2 Preferences and Constraints

All the rational investors choose holdings of the risky asset and of the riskless bond at each date to maximize their expected utility of terminal wealth at $t = 3$. They also have identical preferences that exhibit constant absolute risk aversion. The utility function is

$$u(\tilde{W}_{j3}) = -exp\ (-a\tilde{W}_{j3}), \tag{4.1}$$

where $\tilde{W}_{j3}$ stands for the terminal wealth at $t = 3$ of the informed trader if $j = i$ or of the uninformed if $j = u$. a is the Arrow-Pratt coefficient of absolute risk aversion.

The rational trader's budget constraint at each date can be expressed as follows.

$$\begin{aligned} b_{jt} + z_{jt}p_t &\leq b_{j,t-1} + z_{j,t-1}p_t, \text{ and} \\ b_{j0} + z_{j0}p_0 &\leq \bar{b} + \bar{z}p_0, \quad \text{where } t \in \{1, 2\},\ j \in \{i, u\}. \end{aligned} \tag{4.2}$$

p_t is the price of the risky asset at time t. The riskless and the risky asset holdings of class j trader are b_{jt} and z_{jt} respectively. As a general principle, the first subscript represents the status of the trader, while the second subscript denotes the time period. $\bar{b}$ and $\bar{z}$ are the endowments of riskless and risky asset to each trader at the beginning of $t = 0$. Note also that the unit price of a riskless bond is normalized to one at all times since interest rate is assumed to be zero.

4.2.3 Priors and Information

The random variables of interest to us are: (1) the random dividend payment $\tilde{d}$ of the risky asset at $t = 3$; (2) the random demands $\tilde{z}_{n1}$ and $\tilde{z}_{n2}$ at $t = 1$ and $t = 2$ respectively from the noise traders measured on a per rational trader basis; (3) the y component of the random variable $\tilde{z}_{n2}$ observed by the informed traders. The beliefs of all rational traders can be characterized if we specify a joint distribution of $(\tilde{d}, \tilde{z}_{n1}, \tilde{z}_{n2}, y)^T$. Following the literature, I assume that they are jointly Normal with mean vector $(\mu, 0, 0, 0)^T$ and covariance matrix

$$\begin{bmatrix} \sigma^2 & 0 & 0 & 0 \\ 0 & \sigma_z^2 & 0 & 0 \\ 0 & 0 & \sigma_z^2 & \sigma_y^2 \\ 0 & 0 & \sigma_y^2 & \sigma_y^2 \end{bmatrix}.$$

Due to the property of normal random variables, I here implicitly assume that $\tilde{z}_{n1}$ and $\tilde{z}_{n2}$ are both independent of $\tilde{d}$. $\tilde{z}_{n1}$ and $\tilde{z}_{n2}$ are also mutually independent. $\tilde{z}_{n1}$ is independent of y - the private signal, while the covariance between $\tilde{z}_{n2}$ and y is σ_y^2. This is just a formal way to say that the private signal is informative of noise traders' position at $t = 2$, but not informative of noise traders' position at $t = 1$ or of the random dividend payment $\tilde{d}$.

4.3 Market Equilibrium

The traders maximize their expected utilities at the terminal date by choosing their demand for the risky asset at each date, given their contemporaneous beliefs about the final dividend and the future asset prices. The market clearing condition in turn characterizes the equilibrium price at each date. It is also helpful to clarify the notation a bit. I will consistently use an upper $^\sim$ to indicate that the variable is random from the perspective of the agent, which in turn means those variables without $^\sim$'s are known or are the choice variables.

4.3.1 Equilibrium Asset Price at $t = 2$

At $t = 2$, both informed traders and uninformed traders have the same beliefs about the distribution of future dividend at $t = 3$. For an informed trader, the problem is to choose asset holdings z_{i2} and b_{i2} at $t = 2$ to maximize the terminal expected utility.

$$\begin{aligned} \max_{z_{i2},\, b_{i2}} \quad & E[-e^{-a(z_{i2}\tilde{d}+b_{i2})}|y, p_1, p_2] \\ \text{subject to} \quad & b_{i2} + p_2 z_{i2} \leq W_{i2}. \end{aligned} \tag{4.3}$$

Notice that the only random variable involved is $\tilde{d}$ - the uncertain future dividend of one unit of the risky asset. $W_{i2} \equiv b_{i1} + p_2 z_{i1}$ is the known wealth of an informed trader at date 2 given her choices of asset holdings at date 1. The budget constraint will bind since the objective function is increasing in both z_{i2} and b_{i2}. In the following, I will ignore the inequality in all the budget constraints.

Because none of y, p_1, and p_2 is informative about the uncertain dividend $\tilde{d}$ at $t = 3$, the conditional expectation in the objective equals its unconditional counterpart. Employing a well-known result for Normal random variables, we can simplify the original problem to the maximization of the Certainty Equivalent of $\tilde{W}_{i3} \equiv z_{i2}\tilde{d} + b_{i2}$ with only one choice variable z_{i2}.[29]

$$\max_{z_{i2}} \quad CE_{i2} \equiv z_{i2}(\mu - p_2) + W_{i2} - \frac{a}{2} z_{i2}^2 \sigma^2. \tag{4.4}$$

The first order condition then implies that the demand from this informed trader is

$$z_{i2} = \frac{1}{a\sigma^2}(\mu - p_2). \tag{4.5}$$

Since a typical uninformed trader holds the same belief about $\tilde{d}$ as an informed trader, she will face the same maximization problem except for a different budget $W_{u2} \equiv b_{u1} + p_2 z_{u1}$, which is determined by her choices at date 1. The demand for the risky asset is then

$$z_{u2} = \frac{1}{a\sigma^2}(\mu - p_2). \tag{4.6}$$

Notice that it is equal to the informed trader's demand.

The market clearing condition equalizes the supply to the sum of desired holdings by the informed traders, the uninformed traders, and the noise traders, expressed on a per rational trader basis. The supply at $t = 2$ equals initial endowment $\bar{z}$ because there are zero net supplies both at $t = 1$ and $t = 2$. At $t = 2$, the desired holdings of the risky asset by noise traders equal the cumulated demand from $t = 1$ to $t = 2$, which is $z_{n1} + z_{n2}$. The market clearing condition can then be expressed as follows.

$$\bar{z} = \lambda z_{i2} + (1 - \lambda) z_{u2} + z_{n1} + z_{n2}. \tag{4.7}$$

Notice also that $z_{i2} = z_{u2} = \bar{z} - z_{n1} - z_{n2}$. Therefore, the market clearing price is

$$p_2 = \mu - a\sigma^2(\bar{z} - z_{n1} - z_{n2}). \tag{4.8}$$

[29] The property of the normal random variable used here is: if $\tilde{x} \sim N(\mu, \sigma^2)$, then $E[exp(\tilde{x})] = exp(\mu + \frac{1}{2}\sigma^2)$. Therefore, $E[-exp(-a(z_{i2}\tilde{d} + b_{i2}))] = E[-exp(-a(z_{i2}(\tilde{d} - p_2) + W_{i2}))] = -exp[-az_{i2}(\mu - p_2) - aW_{i2} + \frac{1}{2}a^2 z_{i2}^2 \sigma^2] = -exp[-a(z_{i2}(\mu - p_2) + W_{i2} - \frac{1}{2}a^2 z_{i2}^2 \sigma^2)]$. By definition, $E[-exp(-a(z_{i2}\tilde{d} + b_{i2}))] = -exp(-aCE_{i2})$. As a result, CE_{i2} is as expressed.

At this point, it becomes clear what the private information y is. The private information is correlated with the uncertain demand by liquidity traders $\tilde{z}_{n2}$. Observing equation (4.8), this is equivalent to saying that informed investors have superior knowledge about period 2 price $\tilde{p}_2$.

4.3.2 Rational Expectations Equilibrium at $t = 1$

At $t = 1$, again an informed trader chooses z_{i1} and b_{i1} to maximized its expected utility in period 3.

$$\begin{aligned} \max_{z_{i1},\, b_{i1}} \quad & E[-e^{-a[\tilde{z}_{i2}\tilde{d}+\tilde{b}_{i2}]}|y, p_1] \qquad (4.9) \\ \text{subject to} \quad & b_{i1} + p_1 z_{i1} = W_{i1} \\ & \tilde{b}_{i2} + \tilde{p}_2\tilde{z}_{i2} = b_{i1} + \tilde{p}_2 z_{i1} \\ & \tilde{z}_{i2} = \bar{z} - \tilde{z}_{n1} - \tilde{z}_{n2}. \end{aligned}$$

$W_{i1} \equiv b_{i0} + p_1 z_{i0}$ is the known wealth at $t = 1$. The first two constraints require that the budget for the trader is balanced in both period 1 and period 2. The third equation requires that the informed investor's investment choice $\tilde{z}_{i2}$ at $t = 2$ is optimal, which is implied by equation (4.8) and (4.5).

Substituting these constraints into the original problem, the problem becomes

$$\begin{aligned} \max_{z_{i1}} \quad & E[-e^{-a[\tilde{z}_{i2}(\tilde{d}-\tilde{p}_2)+z_{i1}(\tilde{p}_2-p_1)+W_{i1}]}|y, p_1] \\ \text{subject to} \quad & \tilde{z}_{i2} = \bar{z} - \tilde{z}_{n1} - \tilde{z}_{n2}. \qquad (4.10) \end{aligned}$$

This expression provides some intuition about the problem faced by the informed trader. The terminal wealth at $t = 3$ has three components: the initial wealth W_{i1} at $t = 1$, the appreciation/depreciation in value of her portfolio $z_{i1}(\tilde{p}_2 - p_1)$ from $t = 1$ to $t = 2$, and the appreciation/depreciation in value of her portfolio $\tilde{z}_{i2}(\tilde{d} - \tilde{p}_2)$ from $t = 2$ to $t = 3$. Note that the last two components are random. The informed trader should form beliefs about the distribution of the random components given her current signal y and current price p_1.

Similarly, we can simplify the decision problem by the uninformed investor to

$$\begin{aligned} \max_{z_{u1}} \quad & E[-e^{-a[\tilde{z}_{u2}(\tilde{d}-\tilde{p}_2)+z_{u1}(\tilde{p}_2-p_1)+W_{u1}]}|p_1] \\ \text{subject to} \quad & \tilde{z}_{u2} = \bar{z} - \tilde{z}_{n1} - \tilde{z}_{n2}. \qquad (4.11) \end{aligned}$$

Similar explanation applies to the three components of the terminal wealth in the objective function. The difference is that the uninformed can only condition her belief on the observed price p_1. In order to evaluate this expectation, the uninformed trader should form an expectation about the form of the price function and its joint distribution with other random variables.

I will prove the existence of a rational expectations equilibrium price, and then derive its exact form. The proof and derivation parallels the proof of Theorem 1 in Grossman and Stiglitz (1980). The details are in Appendix C.2.1. Grossman and Stiglitz (1980) also contains a discussion to justify the rational expectations equilibrium by the evolutionary learning of traders.

Proposition 4.1. *Under the joint distribution assumptions made in section 4.2.3, there exists a market clearing price in period 1* $p_1 = \mu + \pi_y y + \pi_z z_{n1} + \bar{\pi}\bar{z}$*, where*

$$\begin{array}{lll}
\pi_y = \frac{a\sigma^2}{a^2\sigma^2+\bar{h}_1}[\lambda h_{i1} + (1-\lambda)h_{u1}\frac{\sigma_y^2+(\pi_z/\pi_y)\sigma_z^2}{\sigma_\psi^2}] & & \\
\pi_z = \pi_y(\frac{a^2\sigma^2}{\lambda h_{i1}}+1) & , & \bar{\pi} = -a\sigma^2 \\
h_{i1} = Var_i(\tilde{z}_{i2}|p_1,y)^{-1} = \frac{1}{\sigma_\epsilon^2} & , & h_{u1} = Var_u(\tilde{z}_{u2}|\psi)^{-1} = \frac{1}{2\sigma_z^2 - \frac{\sigma_y^4+(\pi_z/\pi_y)^2\sigma_z^4}{\sigma_\psi^2}} \\
\bar{h}_1 \equiv \lambda h_{i1} + (1-\lambda)h_{u1} & , & \sigma_\psi^2 = \sigma_y^2 + \frac{1}{\lambda^2}(a^2\sigma^2\sigma_\epsilon^2+\lambda)^2\sigma_z^2 \\
\psi \equiv y + \frac{\pi_z}{\pi_y}z_{n1} & , & \tilde{z}_{i2} = \tilde{z}_{u2} = \bar{z} - \tilde{z}_{n1} - \tilde{z}_{n2}.
\end{array}$$

The demand for the risky asset by an informed investor given private signal y *is*

$$z_{i1} = \frac{\mu - p_1}{a\sigma^2} + \frac{h_{i1}}{a^2\sigma^2}\frac{1}{a\sigma^2}\left(\mu - a\sigma^2 m_{i1} - p_1\right). \tag{4.12}$$

The demand for the risky asset by an uninformed investor given price p_1 *(or equivalently* ψ*) is*

$$z_{u1} = \frac{\mu - p_1}{a\sigma^2} + \frac{h_{u1}}{a^2\sigma^2}\frac{1}{a\sigma^2}\left(\mu - a\sigma^2 m_{u1} - p_1\right), \tag{4.13}$$

where $m_{i1} \equiv E_i(\tilde{z}_{i2}|y,p_1) = \bar{z} - y - z_{n1}$ *and* $m_{u1} \equiv E_u(\tilde{z}_{u2}|\psi) = \bar{z} - \frac{\sigma_y^2+(\pi_z/\pi_y)\sigma_z^2}{\sigma_\psi^2}\psi$*.*

Proof. See Appendix C.2.1 for details. □

To understand these results, the meaning of the several components appearing in the equilibrium price need to be clarified. ψ is just a linear transformation of equilibrium price p_1. It contains the same information as p_1. h_{i1} is the inverse of informed trader's posterior variance of $\tilde{z}_{i2}$. h_{u1} is the inverse of uninformed trader's posterior variance of $\tilde{z}_{u2}$. $\bar{h}_1$ is the weighted average of h_{i1} and h_{u1}. m_{i1} is defined as the posterior mean of $\tilde{z}_{i2}$. m_{u1} is defined as the posterior mean of $\tilde{z}_{u2}$ given current price p_1.

Although $\tilde{z}_{i2} = \tilde{z}_{u2}$, the informed and the uninformed investors hold different information about this random variable. The informed has private information about $\tilde{z}_{n2}$, a component of $\tilde{z}_{i2}$. In addition, she can infer from the price p_1 and her private signal y of the realized random demand from the noise trader z_{n1}, knowing the price is a linear function of y and z_{n1}.

On the other hand, the uninformed trader does not observe the realization of $\tilde{z}_{n1}$, nor does she has any private information of $\tilde{z}_{n2}$. She can only form expectations about the distribution of these two random variables by conditioning on the observed price p_1 or equivalently the signal ψ, rationally expecting the price to be a linear function of y and z_{n1}. As a result, her posterior variance of $\tilde{z}_{u2}$ is higher than that of the informed trader.

To provide more intuition for the results in Proposition 4.1, the following corollary is presented to show a special case when no traders are informed at $t = 1$.

Corollary 4.1. *Under the assumptions made in proposition 4.1, if all the traders are uninformed, i.e. $\lambda = 0$, the equilibrium price is*

$$p_1|_{\lambda=0} = \mu - a\sigma^2(\bar{z} - z_{n1}). \tag{4.14}$$

The homogenous demand for the risky asset is

$$z_{u1}|_{\lambda=0} = \frac{\mu - p_1}{a\sigma^2} + \frac{h_z}{a^2\sigma^2}\frac{1}{a\sigma^2}\Big(\mu - a\sigma^2(\bar{z} - z_{n1}) - p_1\Big) = \bar{z} - z_{n1}, \tag{4.15}$$

where $h_z = 1/\sigma_z^2$.

Proof. Detailed in Appendix C.2.2. □

Comparing equation (4.6) and equation (4.15), we can build some intuition about the effect of price uncertainty at $t = 2$. Given the same price at the two dates, i.e. $p_1 = p_2$, the difference between these two expressions appears to be the added component $\frac{h_z}{a^2\sigma^2}\frac{1}{a\sigma^2}\left(\mu - a\sigma^2(\bar{z} - z_{n1}) - p_1\right)$ in $z_{u1}|_{\lambda=0}$. This term is the speculative demand by the traders. However, this component actually equals zero as shown by equation (4.14).

Comparing equation (4.15) and equation (4.13) and equation (4.12), we note two differences. The first difference is the posterior means of the equal demand from the uninformed and informed traders, namely 0 in $z_{u1}|_{\lambda=0}$, m_{i1} and m_{u1} in z_{i1} and z_{u1} respectively. The second difference lies in the coefficient before the parentheses. It is easy to confirm that $\frac{h_{i1}}{a^2\sigma^2} > \frac{h_{u1}}{a^2\sigma^2}$

and $\frac{h_{i1}}{a^2\sigma^2} > \frac{h_z}{a^2\sigma^2}$. Therefore, the informed trader will be more sensitive to price than will the uninformed trader and than any trader when no private information exists. However, we do not know whether $\frac{h_{u1}}{a^2\sigma^2} > \frac{h_z}{a^2\sigma^2}$, so we do not know if uninformed traders would be more or less sensitive to price than a trader would when there is no private information.

The difference between z_{u1} and $z_{u1}|\lambda = 0$ and that between z_{i1} and $z_{u1}|\lambda = 0$ represent the speculative demand of the uninformed trader and that of the informed trader respectively. As the posterior mean of the random demand from noise traders at $t = 2$ becomes higher, the speculative demands by both the informed and uninformed traders increase. As the posterior variance of the random demand of noise traders becomes lower, the speculative demands by both types of traders also increase as long as the terms in the parentheses are positive.[30]

4.3.3 Equilibrium Asset Price at $t = 0$

The problem at $t = 0$ for the traders is to choose z_0 and b_0 to maximize their expected utilities at the terminal date.

$$\begin{aligned} \max_{z_0,\, b_0} \quad & E[-e^{-a[\tilde{z}_{u2}(\tilde{d}-\tilde{p}_2)+\tilde{z}_{u1}|_{\lambda=0}(\tilde{p}_2-\tilde{p}_1)+z_0(\tilde{p}_1-p_0)+\bar{b}+p_0\bar{z}]}] \quad (4.16)\\ \text{subject to} \quad & \tilde{z}_{u2} = \bar{z} - \tilde{z}_{n1} - \tilde{z}_{n2} \\ & \tilde{p}_2 = \mu - a\sigma^2(\bar{z} - \tilde{z}_{n1} - \tilde{z}_{n2}) \\ & \tilde{z}_{u1}|_{\lambda=0} = \bar{z} - \tilde{z}_{n1} \\ & \tilde{p}_1 = \mu - a\sigma^2(\bar{z} - \tilde{z}_{n1}). \end{aligned}$$

Since at this time all the traders have the same priors about the distribution of $\tilde{z}_{n1}$ and $\tilde{z}_{n2}$, they would rationally expect that $\tilde{p}_1 = \mu - a\sigma^2(\bar{z} - \tilde{z}_{n1})$ and $\tilde{p}_2 = \mu - a\sigma^2(\bar{z} - \tilde{z}_{n1} - \tilde{z}_{n2})$.

Proposition 4.2. *At $t = 0$, after all traders are endowed with the risky asset and the riskless bond, the market clearing price is*

$$p_0 = \mu - a\sigma^2\bar{z}. \tag{4.17}$$

The homogeneous demand by a typical trader is

$$z_0 = \frac{\mu - p_0}{a\sigma^2} + \frac{h_z}{a^2\sigma^2}\frac{1}{a\sigma^2}\left(\mu - a\sigma^2\bar{z} - p_0\right) = \bar{z}. \tag{4.18}$$

[30]Strictly speaking, this is not necessarily true because the speculative demand is a complicated function of posterior variances.

Proof. See Appendix C.2.2 for details. □

The price at $t = 0$ has no random component in it. It hence acts as a benchmark for all future trading prices. μ is the expected dividend value at $t = 3$, and $a\sigma^2\bar{z}$ is the risk premium.

4.3.4 An Example

The model's solution is clearly too complicated to have an analytical comparative static. Therefore, I will examine the comparative statics using numerical analysis. The following example represents a starting point of my parametrization of the model. It resembles the parameterizations in Gennotte and Leland (1990).

The first restriction is on noise trading induced price variation. As we see before, if there is no private information, the variance of future price conditional on current price always equal to $a^2\sigma^4\sigma_z^2$ either at $t = 0$ or at $t = 1$. At $t = 2$, there is no future price but only an uncertain dividend payment $\tilde{d}$, which has a variance of σ^2. It is reasonable to assume that $\sigma^2 = a^2\sigma^4\sigma_z^2$ or $\sigma_z^2 = \frac{1}{a^2\sigma^2}$. This requirement rules out the possibility that trading price may be more or less volatile than the final dividend payment. This is also done in Gennotte and Leland (1990).

A key parameter is the quality of the signal received by the informed trader at $t = 1$. I assume that the signal-to-noise ratio is 0.5. Specifically, I assume that $\sigma_y^2 = 0.5\sigma_z^2$, where σ_y^2 is the variance of signal observed and σ_z^2 is the ex-ante variance of noise trader's random demand at $t = 2$. Therefore, the posterior variance of the informed trader is $\sigma_\epsilon^2 = 0.5\sigma_z^2$.

Another key parameter is the fraction of investors who receive the private signal. It is almost impossible to get an estimate of this number. I assume initially that 10 percent of investors receive the private signal about future random demand. The coefficient of absolute risk aversion is set at 4 (see Barsky et al., 1997, for an estimate of the risk aversion coefficient).

For risk averse traders to hold the risky asset, the risky asset is assumed to have an expected per period return of 6% from $t = 0$ to $t = 3$ and the final dividend payment has a standard deviation of 10%. Assuming the model is in a time window of three years, the total expected return is then 19.1%. Note that the total expected return can be expressed by $\frac{a\sigma^2\bar{z}}{\mu - a\sigma^2\bar{z}}$. If we normalize $\mu - a\sigma^2\bar{z}$ or the price at $t = 0$ to one, it can be shown that the variance of the return is equal to the variance of future dividend or $\sigma^2 = 1\%$. Utilizing the restriction on expected return, we get $\bar{z} = 0.191/(a\sigma^2) = 4.7754$.

The model is now fully specified. The rational expectations equilibrium price at $t = 1$ is

$$p_1 = 1 + 0.0069\tilde{y} + 0.0412\tilde{z}_{n1}. \tag{4.19}$$

Correspondingly, the demand of an informed trader $z_{i1} = 1.4854\tilde{y} - 1.0877\tilde{z}_{n1} + 4.7754$ and the demand of an uninformed trader $z_{u1} = -0.165\tilde{y} - 0.9903\tilde{z}_{n1} + 4.7754$. We can verify that the market clearing condition holds under this demand schedule. p_2 or the price at time $t = 2$ can be computed as $1 + 0.04(\tilde{z}_{n1} + \tilde{z}_{n2})$. The homogeneous demand at $t = 2$ is then $4.7754 - \tilde{z}_{n1} - \tilde{z}_{n2}$.

If no one is informed, the price at $t = 1$ is then $1 + 0.04\tilde{z}_{n1}$, and the demand for the risky asset is $4.7754 - \tilde{z}_{n1}$. The price at $t = 2$ continues to be $1 + 0.04(\tilde{z}_{n1} + \tilde{z}_{n2})$, and the demand for the risky asset is still $4.7754 - \tilde{z}_{n1} - \tilde{z}_{n2}$. It is clear that arrival of private information makes both the informed and uninformed change their positions. Recall that $\tilde{z}_{n1}, \tilde{z}_{n2} \sim N(0, \sigma_z^2 = \frac{1}{a^2\sigma^2} = 6.25\%)$, and $y \sim N(0, 0.5\sigma_z^2 = 3.125\%)$. This numerical example will be our starting point in the following numerical analysis.

4.4 Asset Price Volatility

4.4.1 Calculation of Asset Price Volatility

Comparing the three trading prices when there is no private information $p_0 = \mu - a\sigma^2\bar{z}$, $\tilde{p}_1|_{\lambda=0} = \mu - a\sigma^2(\bar{z} - \tilde{z}_{n1})$, and $\tilde{p}_2 = \mu - a\sigma^2(\bar{z} - \tilde{z}_{n1} - \tilde{z}_{n2})$, we can see that the house price follows a random walk because we assume that $\tilde{z}_{n1}$ and $\tilde{z}_{n2}$ are independent. Define $r_1 = \tilde{p}_1 - p_0$ and $r_2 = \tilde{p}_2 - \tilde{p}_1$ as the price increments at the two dates respectively. It is evident that r_1 and r_2 are identically distributed with a constant variance of $a^2\sigma^4\sigma_z^2$.

When some traders receive private information, the volatility of price increment is no longer a constant. The following proposition summarizes the result.

Proposition 4.3. *If no one receives private signal at $t = 1$, the asset prices in from $t = 0$ to $t = 2$ follow a random walk with independently and identically distributed increments. Specifically, $Var(r_1) = Var(r_2) = a^2\sigma^4\sigma_z^2$.*

When there is private information at $t = 1$, the variances of those increments are no longer a constant of $a^2\sigma^4\sigma_z^2$. Specifically, $Var(r_1) = {\pi_y}^2\sigma_y^2 + {\pi_z}^2\sigma_z^2$, and $Var(r_2) = (\bar{\pi} - \pi_y)^2\sigma_y^2 + (\bar{\pi} - \pi_z)^2\sigma_z^2$.

Proof. The derivation is straightforward using the price function in Proposition 4.1 and the joint distribution assumptions made in section 4.2.3. □

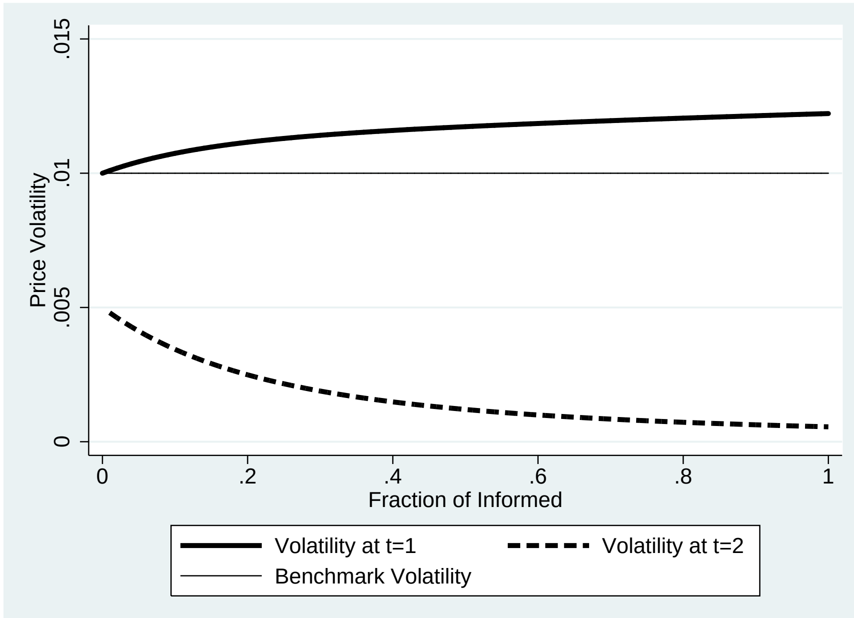

Figure 4.2: Price Volatility and the Fraction of Informed Traders

In the example I presented before, the standard deviation of r_1 is 10.36%, and the standard deviation of r_2 is 5.87%. When there is no private information, the standard deviation of both price increments equals 10%. It seems that private information results in an unbalanced price volatility pattern, which is consistent with the volatility clustering phenomenon observed in housing price series. Intuitively, the speculators destabilize the price at $t = 1$ by speculatively selling/buying the asset, while they stabilize the price at $t = 2$ by neutralizing the partially predicted negative or positive demand shock by the noise traders. Now we are ready to discuss the comparative statics of price volatility with respect to various model parameters.

4.4.2 Comparative Statics of Volatility

The first comparative static exercise involves varying the fraction of informed traders, keeping all other parameters unchanged. We explore the implied change of price increment volatility at $t = 1$ and $t = 2$. Figure 4.2 shows the relation between volatilities and the fraction of traders who receive the private signal. The fraction of informed traders seems to have

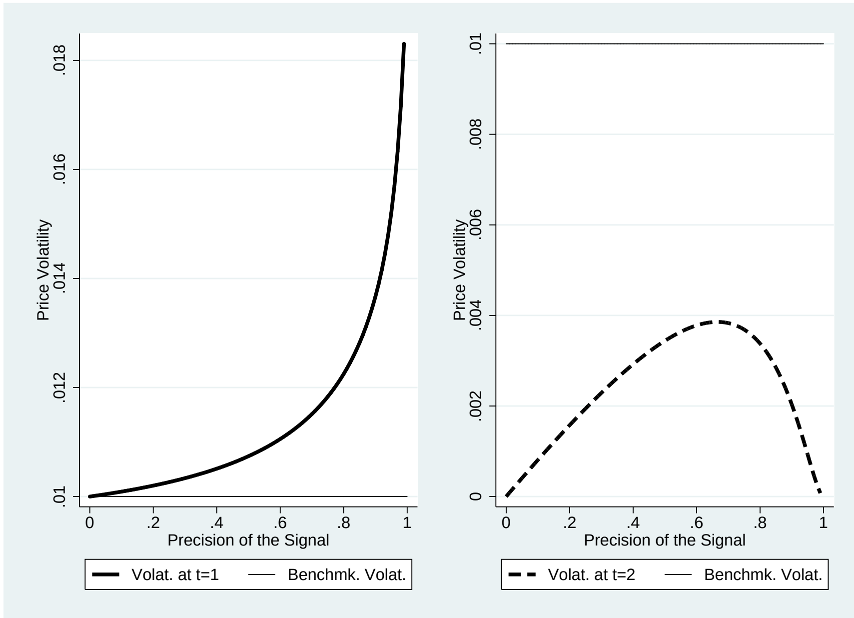

Figure 4.3: Price Volatility and the Precision of the Signal

quite different effect on the volatilities in the two periods. As more traders become informed, the price increment at $t = 1$ becomes more volatile, while the price increment at $t = 2$ becomes less volatile.

Furthermore, the volatility at $t = 2$ is significantly reduced compared with the case when there is no private information. The volatility at $t = 1$ is however increased slightly (up to about 25%). This divergence is mainly due to the speculative activity at $t = 1$ as we discussed at the end of previous section. The intuition is also supported by the finding that the divergence become larger as more traders become informed. Because as more people become informed, they will trade more aggressively than they will when they are uninformed. This will results in an aggravation of price volatility at $t = 1$ but a neutralization of price variability at $t = 2$.

In the next comparative statics exercise, we vary the ratio between the variance of the signal and the variance of the noisy demand by the noise traders at $t = 2$, or $\frac{\sigma_y^2}{\sigma_z^2}$. Recall that the posterior variance is $\sigma_\epsilon^2 = \sigma_z^2 - \sigma_y^2$, so this ratio represents the precision of the signal received by the informed traders. Of course, we keep all other parameters fixed at their original levels.

Figure 4.3 shows the result. The horizontal axis measures the signal to

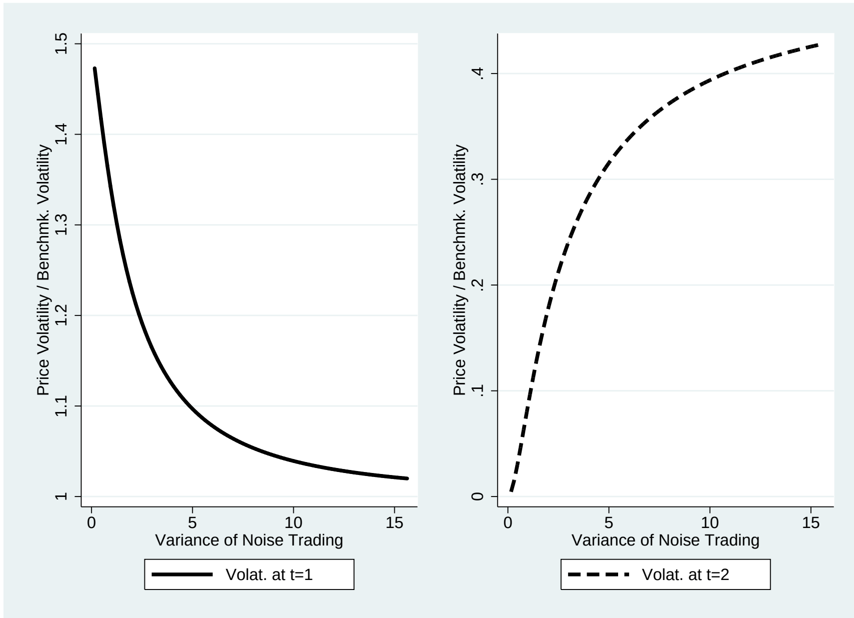

Figure 4.4: Relative Price Volatility and the Variance of Noise Trading

noise ratio mentioned before, taking values from 0.01 to 0.99. Again, the volatility at $t = 1$ increases while the volatility at $t = 2$ decreases significantly compared with the case with no private information. This pattern can be attributed to speculative activity. Intuitively, informed traders would trade more intensively when they get a higher quality signal about future price. This results in more volatile price.

The relationship between period 2 volatility and speculation is more complicated. First, period 2 volatility is significantly reduced once there is speculation. In other words, there is a discontinuous drop in volatility from the no private information case. Second, period 2 volatility initially increases but eventually decreases as the precision of the signal increases. This complicated relationship suggests a complicated effect of signal precision on traders' behavior. I conjecture that speculators may be more likely to make errors in their asset holdings when the signal is imprecise, and they have to reverse this error in period 2. This leads to the initial volatility increase.

Figure 4.4, Figure 4.5, and Figure 4.6 show the comparative statics of volatility with respect to the variance of noise trading, the variance of ter-

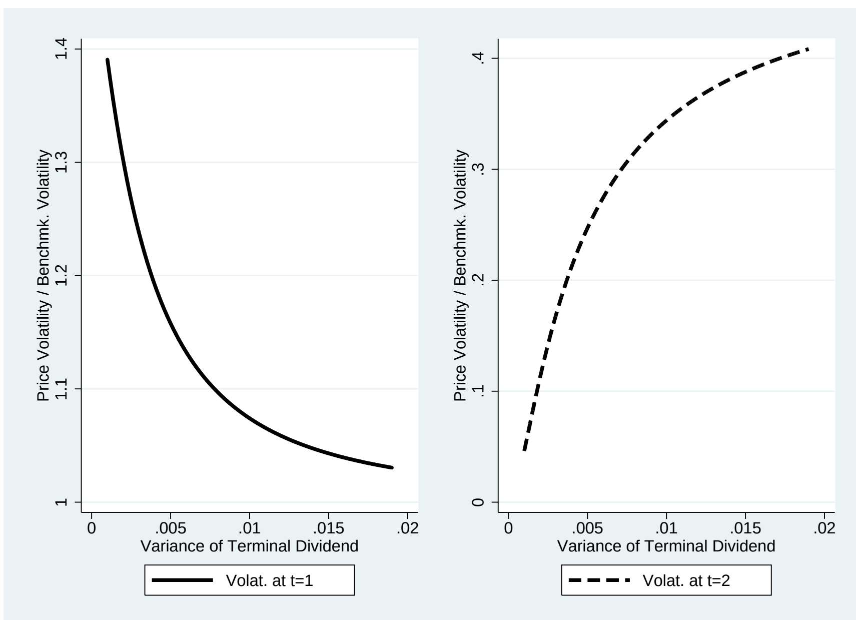

Figure 4.5: Relative Price Volatility and the Variance of Terminal Dividend

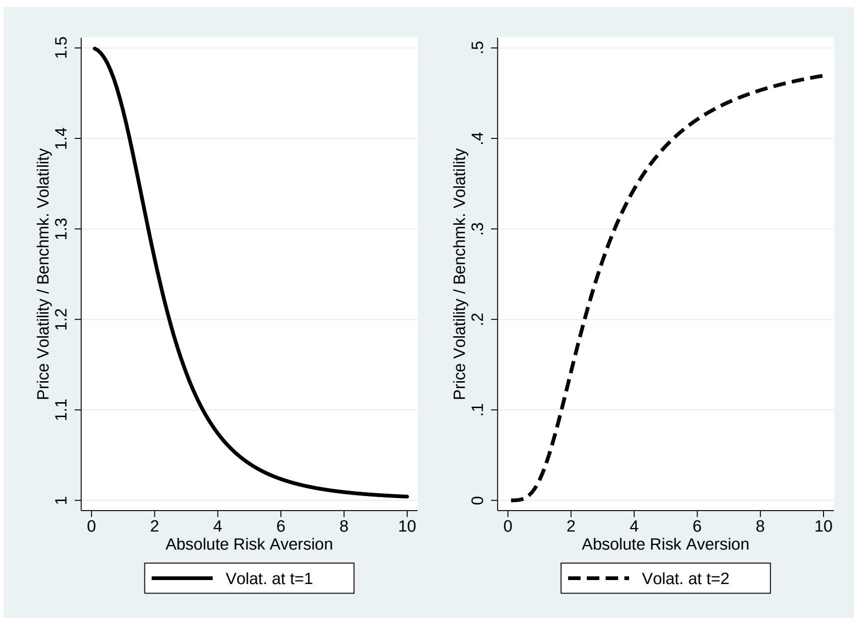

Figure 4.6: Relative Price Volatility and Risk Aversion

minal dividend, and the absolute risk aversion coefficient, respectively. Note that we are now considering the ratio between the observed volatility and the benchmark volatility. I define the benchmark volatility by the volatility of price increment when there is no private information. Because as we change the above three parameters, the volatility under no private information also changes, which makes the comparative statics in terms of levels not informative.

In Figure 4.4, we vary the variance of noise trading from 1/40 of 6.25% to 2.5 times 6.25%, the original variance of noise trading. This means I no longer restrict the model parameters to satisfy the condition that $\sigma_z^2 = \frac{1}{a^2\sigma^2}$. Therefore, price variation may be much smaller or larger than the variation in terminal dividend. Similarly, we vary the variance of terminal dividend between 0.001 to 0.019 in Figure 4.5 while keeping other parameters constant. The above restriction is also relaxed. Figure 4.6 presents the result when we change the absolute risk aversion coefficient from 0.1 to 10.

The comparative statics with respect to those three parameters are fairly similar. The slight increase in price volatility at $t = 1$ and the significant discontinuous drop in price volatility at $t = 2$ are the two recurring results. The similar results are due to the similar underlying economic mechanism. For example, as the variance of noise trading increases, a risk averse speculator would rationally adjust downward their speculative demand, resulting in less volatile period 1 price. At the same time, less speculative activity would has less effect in neutralizing period 2 volatility, resulting in a slight increase in period 2 volatility.

In summary, the comparative statics show a clear pattern about the effect of speculation on price volatility. Even if we assume identically distributed variation in noise trading, we can still get uneven volatility pattern over time. This is mainly due to asymmetric information and speculation. If private information arrives discretely over time, our model can explain the well-documented volatility clustering phenomenon in asset market.

4.5 Return Predictability

4.5.1 Measuring Return Predictability

We've seen before that when there is no private information the house price follows a random walk. Because we assume that $\tilde{z}_{n1}$ and $\tilde{z}_{n2}$ are independent, the two increments r_1 and r_2 are independent and hence have zero correlation.

However, when there is private information, the price increments are no longer independent of each other. They are serially correlated, which can be measured by the 1st order autocorrelation coefficient

$$\rho(1) = \frac{Cov(r_1, r_2)}{\sqrt{Var(r_1)}\sqrt{Var(r_2)}}.$$

The following proposition summarizes the previous discussion and characterize the autocorrelation coefficient when there is private information.

Proposition 4.4. *If no one receives a private signal at $t = 1$, the asset prices in from $t = 0$ to $t = 2$ follow a random walk. Therefore, the autocorrelation coefficient of price increments is zero.*

If there is private information at $t = 1$, the autocorrelation coefficient $\rho_1 \in (0, 1)$ and can be expressed as

$$\rho(1) = \frac{\pi_y(\bar{\pi} - \pi_y)\sigma_y^2 + \pi_z(\bar{\pi} - \pi_z)\sigma_z^2}{[{\pi_y}^2\sigma_y^2 + {\pi_z}^2\sigma_z^2]^{1/2}[(\bar{\pi} - \pi_y)^2\sigma_y^2 + (\bar{\pi} - \pi_z)^2\sigma_z^2]^{1/2}}, \tag{4.20}$$

where π_y, π_z, and $\bar{\pi}$ are defined in Proposition 4.1.

Proof. The derivation is straightforward. Use the derived price function in Proposition 4.1 and employ the joint distribution assumptions made in section 4.2.3 give the result. □

In our example, the calculated first-order autocorrelation coefficient equals 0.0674, which is slightly bigger than zero. Though it is quite small, it still suggests the potential role of speculators in inducing serial correlation in price series. Since we assume throughout that noise traders' new demand are serially uncorrelated, we have shown that serial correlation can happen for purely informational reasons.

4.5.2 Comparative Statics of Autocorrelation Coefficient

The comparative statics of the autocorrelation coefficient with respect to the model parameters are presented in Figure 4.7. These comparative statics provide potential reasons for the observed serial correlation in housing price series.

The range of the underlying parameters is similar to earlier analyses. As the fraction of informed traders and the precision of the signal increase, the autocorrelation coefficient increases monotonically. The former has a marginally decreasing effect, while the latter has a marginally increasing effect. In the extreme cases when the fraction of informed traders is large

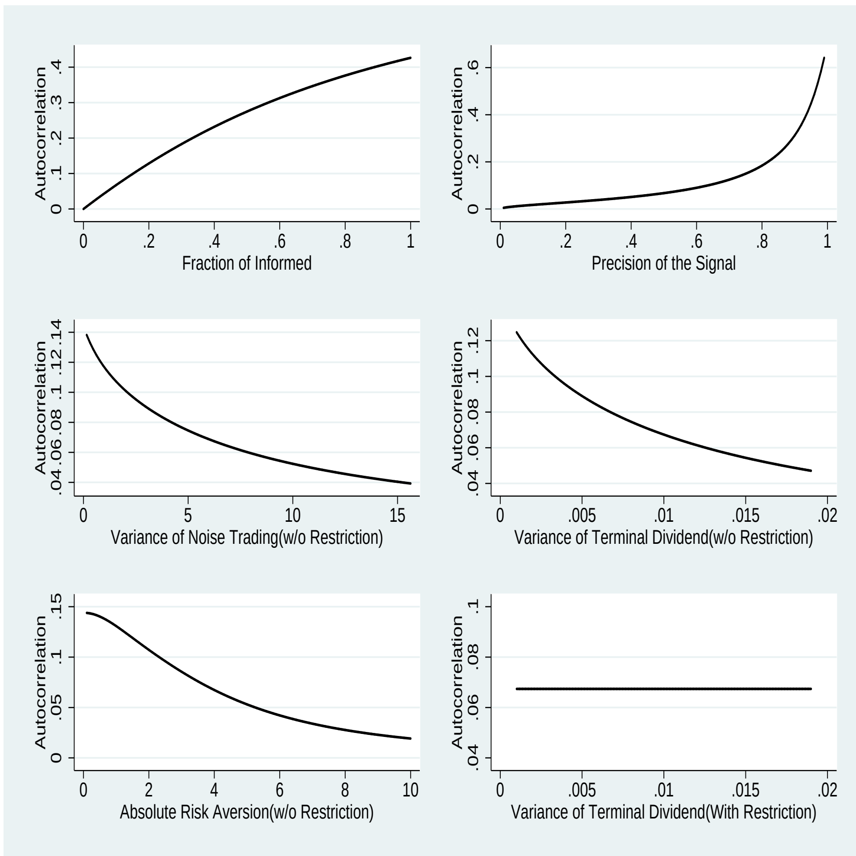

Figure 4.7: Autocorrelation and Model Parameters

or the precision of the signal is very high, the autocorrelation coefficient is around 0.4, which is fairly high.

As the variance of noise trading, the variance of terminal dividend, or the coefficient of absolute risk aversion increases, the autocorrelation coefficient decreases monotonically. Note that we do not impose the constraint that $\sigma_z^2 = \frac{1}{a^2\sigma^2}$ in three of the graphs shown. The results are also intuitive in that all three measures represent the risk associated with the investment in the asset. They would necessarily reduce speculative trading, and hence the induced serial correlation in asset prices.

The last graph at the lower-left corner of Figure 4.7 shows the case when we impose the restriction mentioned above and vary the variance of terminal dividend and the variance of noise trading proportionally, fixing risk aversion and other parameters. Clearly the autocorrelation coefficient does not vary with the variance of the terminal dividend. Though not shown, the autocorrelation also does not vary if we vary the variance of noise trading and risk aversion coefficient proportionally while keeping other parameters constant.

This last graph suggests that serial correlation is determined by the relative magnitude of the price variation due to noise trading and the variation in the terminal payoff, or the ratio between $a^2\sigma^4\sigma_z^2$ and σ^2. Intuitively, this ratio represents the relative importance of noise trading in the price formation process. Paradoxically, if noise trading is relatively unimportant or $a^2\sigma^4\sigma_z^2 < \sigma^2$ or $\sigma_z^2 < \frac{1}{a^2\sigma^2}$, the price tends to be more serially correlated. This is another insight offered by Figure 4.7.

In summary, serial correlation is closely related to speculative activity. Those factors that favor speculative activity: more traders being informed, more precise private signal, less risk averse traders, or less risky terminal payoff and noise trading, tend to increase the serial correlation in returns. In addition, the serial correlation in asset price return is determined by the extent to which asset price volatility is due to noise traders. Paradoxically, when noise trading is relatively unimportant, price tends to be more serially correlated.

4.6 Transaction Volume

4.6.1 Calculation of Transaction Volume

At $t = 0$, there will be no trading if all the traders are endowed with the same amount of risky assets because the desired position for the risky asset is homogeneous (see Proposition 4.2 for details). In other words, there is trading only if the initial endowments are uneven across traders. Since trading, if it exists, is primarily due to the endowment asymmetry, we will not discuss the trading at $t = 0$.

At $t = 1$, there will be trading whether there is private signal or not. The trading volume nevertheless will be different under the two scenarios. Similarly trading will happen at $t = 2$ and the volume is different under the previously mentioned two scenarios. The following proposition summarizes the expected trading volume in each date under the two scenarios.

Proposition 4.5. *When no one receives a private signal at $t = 1$, the expected trading volume at $t = 1$ equals*

$$E(T_1|_{\lambda=0}) = E\Big[\frac{1}{2}(|\tilde{z}_{n1}| + |\tilde{z}_{u1} - \bar{z}|)|_{\lambda=0}\Big] = \Big(\frac{2}{\pi}\Big)^{\frac{1}{2}}\sigma_z. \tag{4.21}$$

The expected trading volume at $t = 2$ equals

$$E(T_2|_{\lambda=0}) = E\Big[\frac{1}{2}(|\tilde{z}_{n2}| + |\tilde{z}_{u2} - \tilde{z}_{u1}|)|_{\lambda=0}\Big] = \Big(\frac{2}{\pi}\Big)^{\frac{1}{2}}\sigma_z. \tag{4.22}$$

In the above, $\tilde{z}_{u1} = \bar{z} - \tilde{z}_{n1}$ and $\tilde{z}_{u2} = \bar{z} - \tilde{z}_{n1} - \tilde{z}_{n2}$.

When there is private information at $t = 1$, the expected trading volume at $t = 1$ equals

$$E(T_1|\lambda) = \Big(\frac{1}{2\pi}\Big)^{\frac{1}{2}}\Big[\sigma_z + \lambda\sqrt{\phi_y^2\sigma_y^2 + \phi_z^2\sigma_z^2} + (1-\lambda)\sqrt{\delta_y^2\sigma_y^2 + \delta_z^2\sigma_z^2}\ \Big]. \tag{4.23}$$

The expected trading volume at $t = 2$ equals

$$\begin{aligned} E(T_2|\lambda) = \Big(\frac{1}{2\pi}\Big)^{\frac{1}{2}}\Big[&\sigma_z + \lambda\sqrt{\sigma_z^2 + \phi_y^2\sigma_y^2 + (1-\phi_z)^2\sigma_z^2 - 2\phi_y\sigma_y^2} \\ &+(1-\lambda)\sqrt{\sigma_z^2 + \delta_y^2\sigma_y^2 + (1-\delta_z)^2\sigma_z^2 - 2\delta_y\sigma_y^2}\ \Big], \end{aligned} \tag{4.24}$$

where

$$\phi_y = \tfrac{\pi_y}{a\sigma^2}(1 + \tfrac{h_{i1}}{a^2\sigma^2}) + \tfrac{h_{i1}}{a^2\sigma^2}, \qquad \delta_y = \tfrac{\pi_y}{a\sigma^2}(1 + \tfrac{h_{u1}}{a^2\sigma^2}) + \frac{h_{u1}(\sigma_y^2 + \frac{\pi_z}{\pi_y}\sigma_z^2)}{a^2\sigma^2\sigma_\psi^2},$$
$$\phi_z = \tfrac{\pi_z}{a\sigma^2}(1 + \tfrac{h_{i1}}{a^2\sigma^2}) + \tfrac{h_{i1}}{a^2\sigma^2}, \qquad \delta_z = \tfrac{\pi_z}{a\sigma^2}(1 + \tfrac{h_{u1}}{a^2\sigma^2}) + \frac{h_{u1}(\sigma_y^2 + \frac{\pi_z}{\pi_y}\sigma_z^2)\frac{\pi_z}{\pi_y}}{a^2\sigma^2\sigma_\psi^2}.$$

Proof. Details are in Appendix C.2.3. □

In our example before, the expected transaction volumes at $t = 1$ and $t = 2$ are 1.9947 if there is no private information. When there is private information, the expected transaction volume at $t = 1$ becomes 2.0432, while the expected transaction volume at $t = 2$ becomes 2.0134. Notice that expected transaction volumes both increase under the asymmetric information case. This is in contrast with the case for price volatility, where the arrival of private information increases the volatility of period 1 but reduces the volatility of period 2.

4.6.2 Comparative Statics of Trading Volume

We have calculated the trading volumes at both trading dates under speculation and asymmetric information. Now we are ready to explore how model parameters affect those numbers and how do they differ from the case under no private information.

The first result is with respect to the fraction of informed shown by Figure 4.8. Again, we vary the fraction from 0 to 1. First, note that if all investors are informed, the trading volumes at both trading dates equal the volumes under no private information. All traders have the same demand for the risky asset, though different from the demand when they are all uninformed. The transaction volume is hence only induced by noise trading. Therefore, the trading volume is the same as the no private information case.

As the fraction of informed traders increases, the transaction volumes at both dates increase initially but decrease eventually to the level under no private information. When only a small fraction of traders are informed, the market price is not very informative of the private signal. As the fraction of informed become larger, the price becomes more informative. However, price informativeness has two opposing effects on the behavior of uninformed traders. First, it increases the willingness for the uninformed to buy and sell the asset since uninformed traders have relatively more precise signals about the future price if the current price becomes more informative. Second, more informative price also reduces the asymmetry between the informed and uninformed. This will lead to less divergence in terms of trading positions.

Figure 4.8 shows the workings of those two factors. Initially, the first effect dominates the second effect and we observe the increase in transaction volume. But the second effect eventually dominates the first by reducing the asymmetry between the informed and uninformed traders. The transaction volume returns to the level when there is no private information.

Another fact shown by Figure 4.8 is the gap between the trading volumes at the two trading dates. Although both are higher than the benchmark trading volume, the volume in period 1 is slightly higher than the volume in period 2. This gap is the largest when both are at their peaks. My conjecture is that the difference is due to the erroneous trading positions

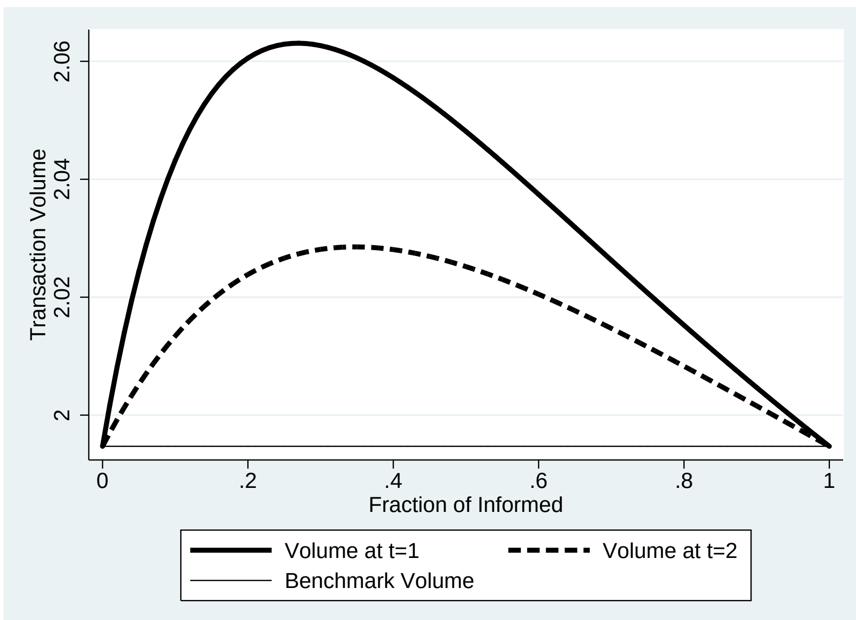

Figure 4.8: Transaction Volume and the Fraction of Informed Traders

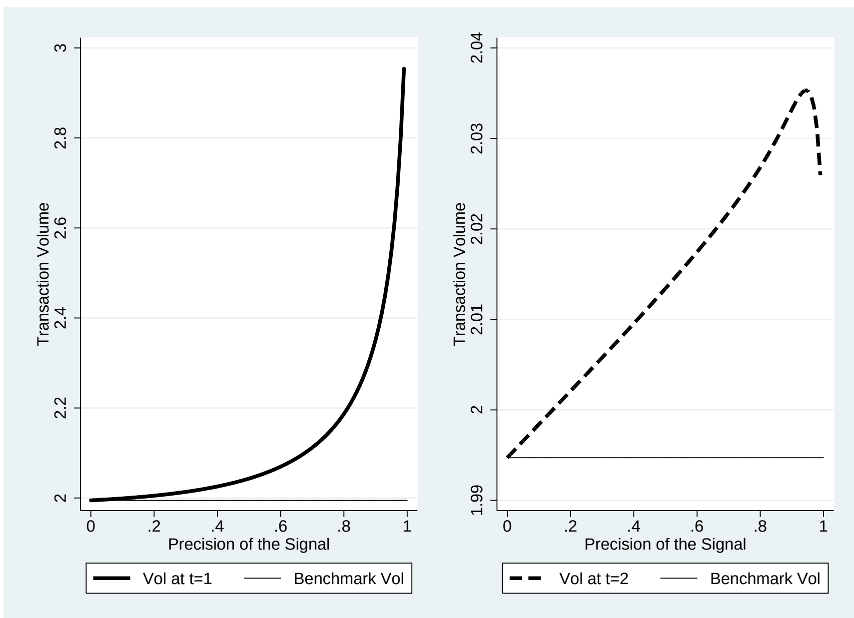

Figure 4.9: Transaction Volume and the Precision of the Signal

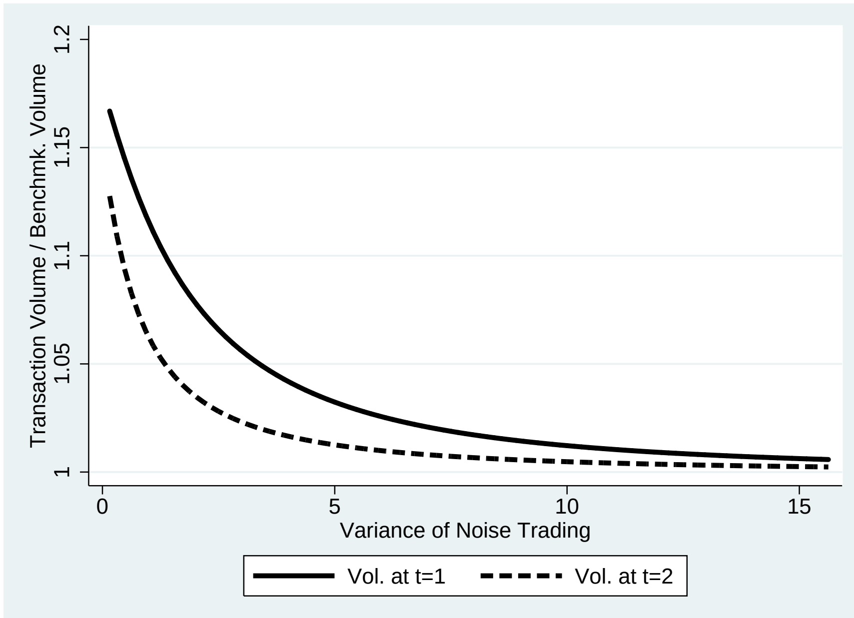

Figure 4.10: Relative Transaction Volume and the Variance of Noise Trading

by the uninformed traders at $t = 1$. Among rational traders, no one makes erroneous judgement in period 2. Therefore, the transaction induced by errors made by uninformed traders no longer exist in period 2. However, the uninformed traders have to reverse their positions at $t = 2$, leading to more transaction volume than the benchmark case.

Figure 4.9 presents the result when we change the precision of the signal. Note that the transaction volume at $t = 1$ is again larger than the the transaction volume at $t = 2$. Both are larger than the benchmark transaction volume under no private information. More precise signal leads to more aggressive speculation on the part of informed traders, which in turn leads to a market price more informative about the private information. In addition, more precise signal also implies less risk associated with the risky asset overall. This then results in more aggressive trading activity by both the informed and uninformed traders. Clearly, the net effect is a monotone increase in transaction volume at $t = 1$. However, the effect on period 2 transaction volume seems to be not monotone, involving an initial increase and an eventual decrease. The intuition for the gap between the trading volumes of the two dates follows our previous discussion of this issue.

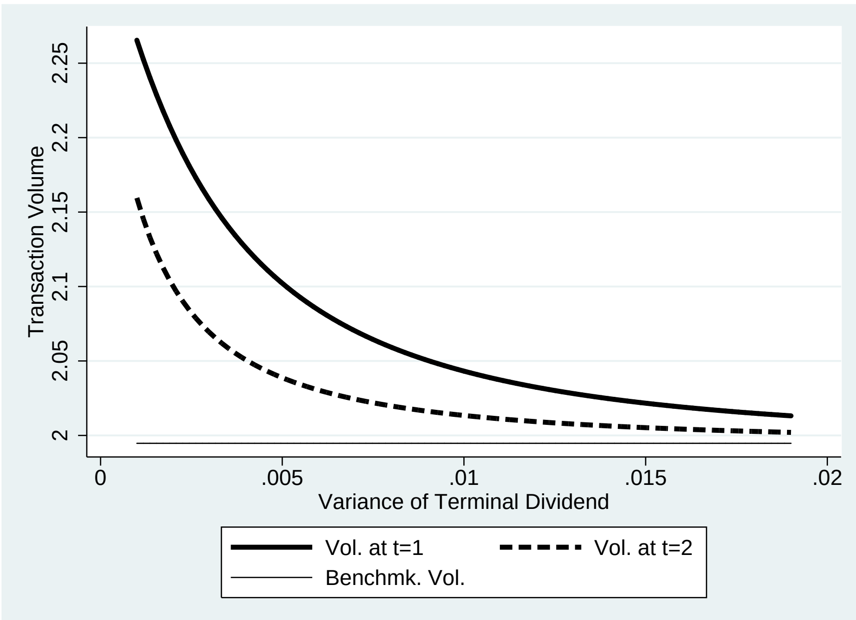

Figure 4.11: Transaction Volume and the Variance of Terminal Dividend

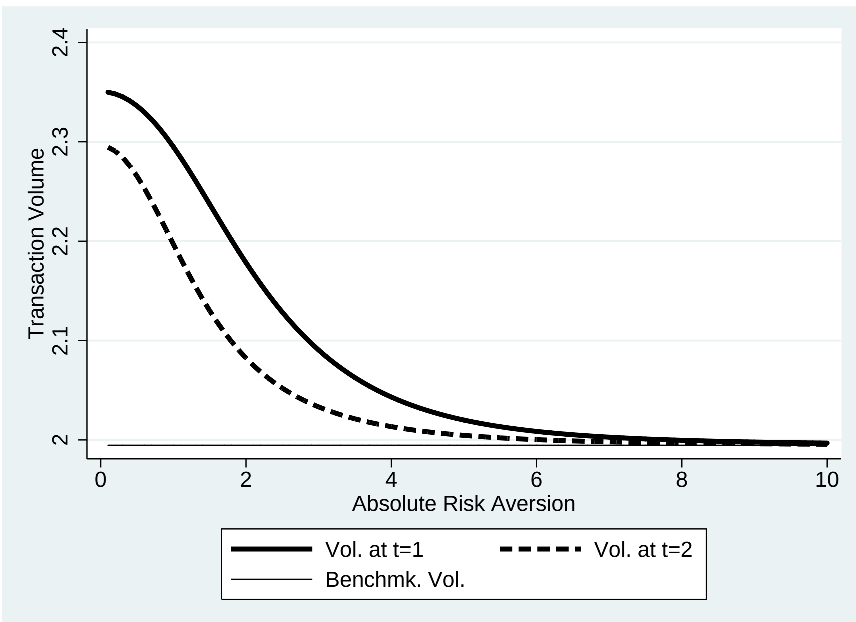

Figure 4.12: Transaction Volume and Risk Aversion

The comparative statics of transaction volume with respect to the variance of noise trading, the variance of terminal dividend, and the risk aversion coefficient are fairly similar. These are shown by Figure 4.10, 4.11, and 4.12 respectively. Note that Figure 4.10 shows the relative transaction volume instead of the level of volume because the benchmark transaction volume also changes when the variance of noise trading changes.

The two major patterns shown before reappear here. All the three graphs show that transaction volumes in both periods are higher than the benchmark volumes. Moreover, the trading volume at $t = 1$ is always higher than the trading volume at $t = 2$. The variance of noise trading and the variance of terminal dividend are both measures of the risk of the asset, while the risk aversion coefficient is a measure of the traders' attitude toward risk. As there is less risk or traders become more tolerant of risk, traders would be more active in speculative trading, resulting in higher transaction volume. The gap between the trading volumes of the two periods can be understood by the same rationale presented before.

In summary of the results shown so far, transaction volume is also closely related to speculation. Whenever the market environment is more favorable to speculation, the transaction volumes tend to increase. There exists, however, a complicated relationship between volume and the fraction of informed and the precision of the private signal. This is because the transaction volume is primarily due to the asymmetric information between the informed and uninformed. Any factor that may affect this asymmetry will also affect the volume, e.g. the informativeness of price.

A special note is warranted at this point. This model can also help understand the price volume correlation in the housing market (see for example Stein, 1995). An extension of the current model by imposing a non-negative constraint on asset holdings may produce such a correlation. If there is a positive demand shock in period 2, the model implies positive expected returns in both period 1 and period 2. At the same time the transaction volume is also higher due to speculation. On the other hand, the model implies negative expected returns in both periods when there is a negative demand shock in period 2. Due to the non-negative constraint, speculation activity is limited and hence the transaction volume will not increase much.

Empirically, we would find a positive correlation between price and transaction volume. In fact, this is purely driven by the unavailability of speculative opportunity when the market condition is bad versus when the market condition is good. This constraint is especially important in the housing market and has long been acknowledged by housing economists (see for example Case and Shiller, 1989). Adding to the literature on price-volume relationship, this model shows that price-volume correlation can happen because of information asymmetry and non-negative constraints on asset holdings.

4.7 Conclusion

In a model of rational expectations equilibrium under asymmetric information, I characterize three summary statics of an asset market: price increment volatility, price increment autocorrelation, and the transaction volume. I show that speculation about noise trading can lead to uneven price volatility over time, serially correlated price increments, and higher transaction volumes.

The driving force is the informed speculation by some traders who can predict future trading flows of noise traders. Whenever the market is more favorable for speculation to happen, the volatility becomes more uneven, and the serial correlation in price increment increases, and the transaction volumes jump up. This model therefore establishes a close connection between speculation and the two striking empirical regularities in the housing price series - serial correlation and volatility clustering.

The model also has a set of comparative statics that can be tested by housing price series. For example, the trading volume is a decreasing function of the variance of asset payoff. Housing price returns become more serially correlated as the variance of asset payoff decreases.

The model, in its simplest form, does not predict a positive correlation between asset price and transaction volume. However, a modified model with non-negative constraint on asset holdings can potentially produce such a correlation. At least, information asymmetry and speculation may be one of the reasons for the existence of price-volume correlation.

Chapter 5

Conclusion and Discussions

To conclude, this dissertation analyzes the complex relationship between language, immigration, and the labor and housing markets in cities. Language barriers lead to labor market segmentation, where wages of workers depend on which labor market they choose to work in. Furthermore, people's tendency to interact within their own language group leads to a positive correlation between housing prices and the homogeneity of ethnic composition in a neighborhood. Last, uncertainty about future immigration/migration flows to a city or a neighborhood, combined with asymmetric knowledge about this information, implies serial correlation and volatility changes in the housing price time-series.

Chapter 2 focuses on wage determination when the labor market is segmented by language barriers. Each worker belongs to either the majority group or the minority group. Each segment is differentiated by the language used at the workplace. The theory predicts two types of wage gaps. The Within-Labor-Market Wage Gap exists between majority workers and minority workers who all belong to the majority market. The Within-Language-Group Wage Gap exists between two types of minority workers, with some working in the majority labor market and others working in the minority labor market. The theory also implies that individuals' wages depend on language group population in a certain way.

My empirical analysis of Canadian Census Public Use Micro-Data on individuals (2001) confirms the above theory predictions. The special feature of the data is the reported language at work, which I use to identify a worker's labor market segment. I employ switching regression methods to identify two types of wage gaps, with self-selection corrected using Heckman two-step procedures. I also use 2SLS to address the concern that population may be endogenous to wage determination.

This theory of labor market segmentation has interesting implications. For example, immigration policies that reduce language barriers may have unexpected effects. The model shows that minority workers in the ethnic enclaves will get lower earnings, while those who work in the majority labor market will benefit. In fact, majority workers in the majority labor market will also benefit from such policies.

Chapter 3 focuses on the effects of ethnic preference on the housing market. If people prefer to interact with those who belong to the same ethnic group, people will have friends primarily from the same ethnic group. In addition, they will choose a location where they can have easy access to their friends. Therefore, housing prices will be higher for locations that have higher quality of social interactions, which is directly related to the ethnic composition of a neighborhood. I characterize a neighborhood choice model in which housing prices and the composition of ethnic groups are endogenous. I also present an example where the average house price can be correlated with the Herfindahl index of ethnic group composition.

The empirical analysis tests a correlation between the average house price and the Herfindahl index of ethnic group composition. I combine housing transaction data in Vancouver with Census data to cross-sectionally test this relationship. I find that the house price and the Herfindahl index of only minority groups, excluding the English speaking group, are positively correlated. As the theory shows, the Herfindahl index in fact is endogenous to the house price equation. To alleviate this concern, I carry out fixed effect regressions and also the conditional difference-in-difference analysis, which allow such endogeneity to exist. I find very similar positive results.

The finding in Chapter 3 is interesting. First, it shows that the preference toward interactions within ethnic group boundary differs across groups. I show that majority individuals may not care about the ethnic composition because they have very diverse social networks. Therefore, it is the homogeneity of the minorities in a neighborhood that predicts higher housing prices. Second, both policy makers and practitioners are interested in finding better predictors of housing prices. They may use the result in this chapter to formulate better zoning regulations or make profits by choosing the right place to develop a project.

Chapter 4 analyzes the properties of housing prices when a city or a neighborhood is subject to an uncertain inflow/outflow of migrants. In the model, the price acts not only as an instrument to allocate resources but also as a source of information about the future. Some people have better knowledge of the future migration flows, while others rely primarily on the price for predictions about the future.

I find that prices become serially correlated and that price volatility changes over time. These are two puzzling phenomena documented in the housing literature. I also provide a rich set of comparative statics, which future researchers may want to test.

As for future research, there are at least three promising avenues. First, we need to provide a microfoundation for people's preference toward social interactions. A microfoundation will provide more concrete empirical predictions to match with the data. Specifically, we need to clarify the benefits and costs involved in people's choices of social interactions so that we can relate them to the observed data. I suspect a framework similar to that of labor supply may be fruitful.

Second, it may be interesting to model how economic forces affect people's language choices. In other words, how does a multicultural society evolve over time? Additionally, does social welfare improve along the process? These are difficult questions, but it certainly is of interest to policy makers and individuals.

Finally, information asymmetry is prevalent in the real estate market. Yet there is still a lot of work to be done by applying the tools from information economics to the analysis of housing price determination, real estate development patterns, and the role of brokers in the market.

Bibliography

Alesina, A. and E. LaFerrara (2002). Who trusts others? *Journal of Public Economics 85*(2), 207–234.

Alesina, A. and E. LaFerrara (2005). Ethnic diversity and economic performance. *Journal of Economic Literature 43*(3), 762–800.

Barsky, R. B., F. T. Juster, M. S. Kimball, and M. D. Shapiro (1997). Preferenc parameters and behavioral heterogeneity: An experiemental approach in the health and retirement study. *Quarterly Journal of Economics 112*(2), 537–579.

Baumol, W. J. (1957). Speculation, profitability, and stability. *Review of Economics and Statistics 39*(3), 263–271.

Bayer, P., R. McMillan, and K. Rueben (2005). An equilibrium model of sorting in an urban housing market. working paper.

Bleakley, H. and A. Chin (2004). Language skills and earnings: Evidence from childhood immigrants. *Review of Economics and Statistics 86*(2), 481–496.

Capozza, D. R. and R. W. Helsley (1989). The fundamentals of land prices and urban growth. *Journal of Urban Economics 26*(3), 295–306.

Capozza, D. R., P. H. Hendershott, and C. Mack (2004). An anatomy of price dynamics in illiquid markets: Analysis and evidence from local housing markets. *Real Estate Economics 32*(1), 1–32.

Card, D. (1999). The causal effect of education on earnings. In O. C. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics*, Volume 3A, Amsterdam, pp. 1801–1863. North-Holland.

Card, D. (2001). Immigrant inflows, native outflows and the local labor market impacts of higher immigration. *Journal of Labor Economics 19*(1), 22–61.

Case, K. E. and R. J. Shiller (1989). The efficiency of the market for single-family homes. *American Economic Review 79*(1), 125–137.

Chiswick, B. R. and P. W. Miller (1992). Language in the immigrant labor market. In B. R. Chiswick (Ed.), *Immigration, Language, and Ethnicity*, Washington, D.C., pp. 229–287. The AEI Press.

Chiswick, B. R. and P. W. Miller (1995). The endogeneity between language and earnings: International analyses. *Journal of Labor Economics 13*(2), 246–288.

Chiswick, B. R. and P. W. Miller (2004). Linguistic distance: A quantitative measure of the distance between english and other languages. *IZA Discussion Paper No. 1246*.

Christensen, P. O. and G. A. Feltham (2002). *Economics of Accounting*, Volume I: Information in Markets. Hingham, MA USA: Kluwer Academic Publishers.

CIC (2001). Towards a more balanced geographic distribution of immigrants. *A special study by Citizenship and Immigriation Canada(CIC): Strategic Planning, Policy and Research*.

Dickens, W. T. and L. F. Katz (1987). Industry wage differences and theories of wage determination. *NBER Working Paper 2271*.

Dickens, W. T. and K. Lang (1985). A test of dual labor market theory. *American Economic Review 75*(4), 792–805.

Dickens, W. T. and K. Lang (1992). Labor market segmentation thoery: Reconsidering the evidence. *NBER working paer No. 4087*.

Farrell, M. J. (1966). Profitable speculation. *Economica 33*(130), 183–193.

Gennotte, G. and H. Leland (1990). Market liquidity, hedging, and crashes. *American Economic Review 80*(5), 999–1021.

Glaeser, E. L. and D. C. Mare (2001). Cities and skills. *Journal of Labor Economics 19*(2), 316–342.

Grossman, S. J. and J. E. Stiglitz (1980). On the impossibility of informationally efficient market. *American Economic Review 70*(3), 393–408.

Hart, O. D. and D. M. Kreps (1986). Price destabilizing speculation. *Journal of Political Economy 94*(5), 927–951.

Heckman, J., H. Ichimura, J. Smith, and P. Todd (1998). Characterizing selection bias using experimental data. *Econometrica 66*(5), 1017–1098.

Helsley, R. W. and W. C. Strange (1990). Matching and agglomeration economies in a system of cities. *Regional Science and Urban Economics 20*(2), 189–212.

Jacobs, J. (1961). *The Death and Life of Great American Cities* (Modern Library, 1993 ed.). New York, NY: Random House.

Jacobs, J. (1969). *The Economy of Cities*. New York, NY: Random House.

Lang, K. (1986). A language theory of discrimination. *Quarterly Journal of Economics 101*(2), 363–382.

Lee, L. (1982). Some approaches to the correction of selectivity bias. *Review of Economic Studies 49*(3), 355–372.

Mankiw, N. G. and D. N. Weil (1989). The baby boom, the baby bust, and the housing market. *Regional Science and Urban Economics 19*(2), 235–258.

Mas-Colell, A., M. D. Whinston, and J. R. Green (1995). *Microeconomic Theory*. New York, NY: Oxford University Press.

Miquel, R. (2003). Identification of effects of dynamic treatments with a difference-in-differences approach. *Discussion Paper 2003-06*. Department of Economics, University of St. Gallen.

Miyao, T. (1978a). Dynamic instability of a mixed city in the presence of neighborhood externalities. *American Economic Review 68*(3), 454–463.

Miyao, T. (1978b). A probabilistic model of location choice with neighborhood effects. *Journal of Economic Theory 19*(2), 347–358.

Rosenbaum, P. and D. Rubin (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika 70*, 41–55.

Schelling, T. C. (1969). Models of segregation. *American Economic Review 59*(2), 488–493.

Schelling, T. C. (1971). Dynamic models of segregation. *Journal of Mathematical Sociology 1*, 143–186.

Stein, J. C. (1995). Prices and trading volume in the housing market: A model with down-payment effects. *Quarterly Journal of Economics 110*(2), 379–406.

Sutton, J. (1986). Non-cooperative bargaining theory: An introduction. *Review of Economic Studies 53*(5), 709–724.

Telser, L. G. (1959). A theory of speculation relating profitability and stability. *Review of Economics and Statistics 41*(3), 295–301.

Wardhaugh, R. (2005). *An Introduction to Sociolinguistics* (5 ed.). Blackwell Publishing Limited.

Wong, M. (2007). Estimating ingroup preferences using ethnic housing quotas in singapore. *Working Paper*.

Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.

Yinger, J. (1976). Racial prejudice and racial residential segregation in an urban model. *Journal of Urban Economics 3*(4), 383–396.

Appendix A

Appendix for Chapter 2

A.1 Derivations

A.1.1 The Derivation of Matching Quality

We begin with the matching quality of majority group workers. First, note that $|x-y|$ is a uniform random variable over the interval $[y, y+1/2]$ on the unit circle. Its probability density is 2. Therefore, the probability that a worker with characteristic x belongs to the market area of a firm with address y is

$$Pr(|x-y| < 1/2m_l) = \int_0^{1/2m_l} 2d\mu = 1/m_l. \tag{A.1}$$

The event that a worker x is employed by firm y is a Bernoulli random variable. The number of majority workers employed by the firm $\Omega_l^d(y)$ therefore follows a binomial distribution with parameters n_l^d and $1/m_l$. The expected employment is hence $E[\Omega_l^d(y)] = \frac{n_l^d}{m_l}$.

The expected distance on the unit circle between a firm's employee and the firm can be expressed as $E[|x-y| : x \in (y - \frac{1}{2m_l}, y + \frac{1}{2m_l})]$. The probability density distribution function of $|x-y|$ conditional on the worker belongs to the firm's market area is

$$f\left[|x-y| : x \in (y-\frac{1}{2m_l}, y+\frac{1}{2m_l})\right] = \begin{cases} \frac{2}{Pr[x \in (y-\frac{1}{2m_l}, y+\frac{1}{2m_l})]} & if\ x \in (y - \frac{1}{2m_l}, y + \frac{1}{2m_l}) \\ 0 & if\ x < y - \frac{1}{2m_l}\ or\ x > y + \frac{1}{2m_l} \end{cases}$$

The denominator is $1/m_l$ (equation (A.1)) in the first case. Therefore, the conditional expectation of $|x-y|$ equals

$$E[|x-y| : x \in (y - \frac{1}{2m_l}, y + \frac{1}{2m_l})] = \int_0^{1/2m_l} 2m_l \mu d\mu = 1/4m_l. \tag{A.2}$$

The expected distance decreases as m_l increases. The expected distance between a minority worker and the firm can be derived analogously. It is also $1/4m_l$.

A.1.2 Proof of Lemma 2.2

First, I derive the expected profit in the majority labor market $E(\pi_{l_d})$. We know that $\pi_{l_d} = q_{l_d} - \kappa_{l_d}$. The combination of equation (2.4), (2.5), and (2.6) gives

$$\begin{aligned}\pi_{l_d}\Big(y, \Omega^d_{l_d}(y), \Omega^m_{l_d}(y)\Big) &= \tfrac{1}{2}\Big\{ \textstyle\sum_{i \in \Omega^d_{l_d}(y)} \Big(\alpha^d + \xi_i - \beta^d_{l_d}|x_i - y|\Big) \\ &\quad + \textstyle\sum_{i \in \Omega^m_{l_d}(y)} \Big(\alpha^m + \xi_i - \beta^m_{l_d}|x_i - y|\Big)\Big\} - C_{l_d}.\end{aligned}$$

The expected revenue earned from the set of majority group workers $\Omega^d_{l_d}$ can be expressed as

$$\begin{aligned}&E[\textstyle\sum_{i \in \Omega^d_{l_d}(y)} (\alpha^d + \xi_i - \beta^d_{l_d}|x_i - y|)] \\ &= E[\textstyle\sum_{i=1}^{\Omega^d_{l_d}(y)} (\alpha^d + \xi_i - \beta^d_{l_d}|x_i - y|) : i \in \Omega^d_{l_d}(y)] \\ &= E[\textstyle\sum_{i=1}^{\Omega^d_{l_d}(y)} (\alpha^d + \xi_i - \beta^d_{l_d}|x_i - y|) : |x_i - y| < \frac{1}{2m_{l_d}}] \\ &= E[\textstyle\sum_{i=1}^{\Omega^d_{l_d}(y)} (\alpha^d + \xi_i)] - E[\textstyle\sum_{i=1}^{\Omega^d_{l_d}(y)} \beta^d_{l_d}|x_i - y| : |x_i - y| < \frac{1}{2m_{l_d}}] \\ &= E[\Omega^d_{l_d}(y)]\alpha^d - E[\Omega^d_{l_d}(y)]E[|x_i - y| : |x_i - y| < \frac{1}{2m_{l_d}}]\beta^d_{l_d} \\ &= (n^d/m_{l_d})(\alpha^d - \beta^d_{l_d}/(4m_{l_d})).\end{aligned}$$

The first and second equalities are re-expressions of the expected revenue. The third equality uses the assumption that x_i is independent of both ξ_i and ψ_i, which implies the independence between $\Omega^d_{l_d}$ and x_i, and also ξ_i and x_i. The next equality employs the independence between $\Omega^d_{l_d}$ and x_i and assumption about ξ_i. The last equality employs the results in section A.1.1.

The expected revenue earned from the set of minority workers $\Omega^m_{l_d}$ is complicated.

$$\begin{aligned}&E[\textstyle\sum_{i \in \Omega^m_{l_d}(y)} (\alpha^m + \xi_i - \beta^m_{l_d}|x_i - y|)] \\ &= E[\textstyle\sum_{i=1}^{\Omega^m_{l_d}(y)} (\alpha^m + \xi_i - \beta^m_{l_d}|x_i - y|) : i \in \Omega^m_{l_d}(y)] \\ &= E[\textstyle\sum_{i=1}^{\Omega^m_{l_d}(y)} (\alpha^m + \xi_i - \beta^m_{l_d}|x_i - y|) : \psi_i < \mu, |x_i - y| < \frac{1}{2m_{l_d}}] \\ &= E[\textstyle\sum_{i=1}^{\Omega^m_{l_d}(y)} (\alpha^m + \xi_i) : \psi_i < \mu] - E[\textstyle\sum_{i=1}^{\Omega^m_{l_d}(y)} \beta^m_{l_d}|x_i - y| : \psi_i < \mu, |x_i - y| < \frac{1}{2m_{l_d}}] \\ &= E[\Omega^m_{l_d}(y)]\alpha^m + E[\Omega^m_{l_d}(y)]E[\xi_i : \psi_i < \mu] - E[\Omega^m_{l_d}(y)]E[|x_i - y| : |x_i - y| < \frac{1}{2m_{l_d}}]\beta^m_{l_d} \\ &= (n^m \cdot \Phi(\mu/\sigma)/m_{l_d})(\alpha^m - \rho s \tfrac{\phi(\mu/\sigma)}{\Phi(\mu/\sigma)} - \beta^m_{l_d}/(4m_{l_d})).\end{aligned}$$

The difference lies in the second equality, where I add $\psi_i < \mu$ where $\mu = (\beta^m_{l_m}/8m_{l_m}) - (\beta^m_{l_d}/8m_{l_d}) - \theta^m_{l_d}$. Only those workers whose ψ's satisfy the inequality enter the majority labor market. The third equality again uses the independence of x_i with both ψ_i and ξ_i. The next equality uses the independence between $\Omega^m_{l_d}$ and ψ_i. Strictly speaking, this is only approximately true as n^m becomes large enough.[31] The last equality uses joint normality of ξ_i and ψ_i, and $\frac{\phi(\mu/\sigma)}{\Phi(\mu/\sigma)}$ is the inverse Mill's ratio. The last equality uses results in section A.1.1 as well.

Substituting the above two terms into the expected profit expression gives the results of the lemma. Analogously, I can derive the expected profit in the minority labor market.

A.1.3 Proof of Proposition 2.1

Under assumptions made in Proposition 2.1, the two equilibrium conditions are

$$E(\pi_{l_d}) = \frac{n^d}{2m_{l_d}}(\alpha^d - \frac{\beta^d_{l_d}}{4m_{l_d}}) + \frac{n^m \cdot \Phi(\mu/\sigma)}{2m_{l_d}}[\alpha^m - \frac{\beta^m_{l_d}}{4m_{l_d}}] - C_{l_d} = 0, \text{ and}$$

$$E(\pi_{l_m}) = \frac{n^m \cdot (1 - \Phi(\mu/\sigma))}{2m_{l_m}}[\alpha^m - \frac{\beta^m_{l_m}}{4m_{l_m}}] - C_{l_m} = 0.$$

Note $\mu = (\beta^m_{l_m}/8m_{l_m}) - (\beta^m_{l_d}/8m_{l_d}) - \theta^m_{l_d}$. The selection correction terms disappear because $\rho = 0$. The Jacobian J with respect to m_{l_d} and m_{l_m} of the two equations is

$$\begin{bmatrix} n^d \frac{\partial F_1}{\partial m_{l_d}} + n^m \frac{\partial \Phi}{\partial m_{l_d}} F_2 + n^m \Phi \frac{\partial F_2}{\partial m_{l_d}} & n^m \frac{\partial \Phi}{\partial m_{l_m}} F_2 \\ -n^m F_3 \frac{\partial \Phi}{\partial m_{l_d}} & -n^m F_3 \frac{\partial \Phi}{\partial m_{l_m}} + n^m (1 - \Phi) \frac{\partial F_3}{\partial m_{l_m}} \end{bmatrix}.$$

Here $F_1 = \frac{1}{2m_{l_d}}(\alpha^d - \frac{\beta^d_{l_d}}{4m_{l_d}})$, $F_2 = \frac{1}{2m_{l_d}}(\alpha^m - \frac{\beta^m_{l_d}}{4m_{l_d}})$, and $F_3 = \frac{1}{2m_{l_m}}(\alpha^m - \frac{\beta^m_{l_m}}{4m_{l_m}})$. It is straightforward to verify that $\frac{\partial E[\pi_{l_d}]}{\partial m_{l_m}} < 0$ and $\frac{\partial E[\pi_{l_m}]}{\partial m_{l_d}} < 0$. Because J is assumed to be negative definite, $\frac{\partial E[\pi_{l_d}]}{\partial m_{l_d}} < 0$ and $\frac{\partial E[\pi_{l_m}]}{\partial m_{l_d}} < 0$.

$E(\pi_{l_d})$ and $E(\pi_{l_m})$ are continuously differentiable with respect to m_{l_d} and m_{l_m}. We assume that J is negative definite, so $|J| > 0$. The comparative statics is hence well-defined due to the Implicit Function Theorem. Take total differential of the equation system, we get

$$\begin{aligned} \frac{\partial E[\pi_{l_d}]}{\partial m_{l_d}} dm_{l_d} + \frac{\partial E[\pi_{l_d}]}{\partial m_{l_m}} dm_{l_m} &= -\Big(\frac{\partial E[\pi_{l_d}]}{\partial n_{l_d}} dn_{l_d} + \cdots + \frac{\partial E[\pi_{l_d}]}{\partial \sigma} d\sigma \Big) \\ \frac{\partial E[\pi_{l_m}]}{\partial m_{l_d}} dm_{l_d} + \frac{\partial E[\pi_{l_m}]}{\partial m_{l_m}} dm_{l_m} &= -\Big(\frac{\partial E[\pi_{l_m}]}{\partial n_{l_d}} dn_{l_d} + \cdots + \frac{\partial E[\pi_{l_m}]}{\partial \sigma} d\sigma \Big), \end{aligned}$$

[31] Because ψ_i's are independent across workers, the random variable $n^m_{l_d}$ is approximately independent of a particular ψ_i as n^m increases.

where I omit several exogenous variables to simplify the expression. Let all the differentials be zeros except dn_{l_d} and then divide both sides of the equation system by dn_{l_d}. I get

$$\begin{bmatrix} \frac{\partial E[\pi_{l_d}]}{\partial m_{l_d}} & \frac{\partial E[\pi_{l_d}]}{\partial m_{l_m}} \\ \\ \frac{\partial E[\pi_{l_m}]}{\partial m_{l_d}} & \frac{\partial E[\pi_{l_m}]}{\partial m_{l_m}} \end{bmatrix} \begin{bmatrix} \frac{\partial m_{l_d}}{\partial n_{l_d}} \\ \\ \frac{\partial m_{l_m}}{\partial n_{l_d}} \end{bmatrix} = \begin{bmatrix} -\frac{\partial E[\pi_{l_d}]}{\partial n_{l_d}} \\ \\ -\frac{\partial E[\pi_{l_m}]}{\partial n_{l_d}} \end{bmatrix}$$

By Cramer's rule, the solution of this system can be written

$$\frac{\partial m_{l_d}}{\partial n_{l_d}} = \frac{|J_1|}{|J|}, \text{ and } \frac{\partial m_{l_m}}{\partial n_{l_d}} = \frac{|J_2|}{|J|}, \tag{A.3}$$

where $|.|$ denotes the determinant of the matrix, J is the Jacobian of the equation system, and J_i is the new matrix formed by substituting the right-hand side column vector to the ith column of the Jacobian.

It is easy to calculate that $-\frac{\partial E[\pi_{l_d}]}{\partial n_{l_d}} = -F_1$ and $-\frac{\partial E[\pi_{l_m}]}{\partial n_{l_d}} = 0$. We also know that all the terms in the Jacobian are negative. Using equation (A.3), it is straightforward to show that $\frac{\partial m_{l_d}}{\partial n_{l_d}} > 0$ and $\frac{\partial m_{l_m}}{\partial n_{l_d}} < 0$. The comparative statics with respect to other exogenous variables α^d, C_{l_m}, $\beta^m_{l_m}$, $\theta^m_{l_d}$, C_{l_d}, $\beta^d_{l_d}$, $\beta^m_{l_d}$, n^m, α^m, and σ can be derived following similar steps.

A.2 A Generalized Model of Labor Market Segmentation

This section contains a generalization of the basic model to multiple language groups and multiple labor markets. I continue to use superscript $k = 1, ..., K$ to represent a worker's language group, where "1" denotes the majority group and K is the number of language groups. All other groups are minorities. The population of a language group is denoted n^k.

I use subscript $l = 1, ..., L$ to index the labor market, where "1" represents the majority labor market and L is the number of labor markets. The number of workers who belong to language group k and work in labor market l is denoted n^k_l. Note that $K = L$ must hold.

Assumption A.1. (1) Majority group workers work only in the majority market;
(2) Minority workers can choose only between the majority labor market and their own language groups' labor market.
(3) The conditions in Assumption 2.1 hold.

Under Assumption A.1, the choice for a minority worker of language group k is between labor market $l = 1$ and labor market $l = k$. The expected net wage can be expressed as

$$E\left(U_{l,i}^k\right) = \begin{cases} \frac{1}{2}[\alpha^k + \xi_i - \frac{\beta_1^k}{4m_1}] - \theta_1^k - \psi_i & \text{if } l = 1, \text{ and } k = 2, ..., K \\ \\ \frac{1}{2}[\alpha^k + \xi_i - \frac{\beta_k^k}{4m_k}] & \text{if } l = k, \text{ and } k = 2, ..., K \end{cases}. \quad \text{(A.4)}$$

The minority worker chooses the majority labor market if and only if $(\beta_k^k/8m_k) - (\beta_1^k/8m_1) - \theta_1^k - \psi_i > 0$.

The number of minority workers of language group $k \geq 2$ in the majority labor market $n_{l_m}^k$ is a binomial random variable. Its expectation is $n^k\Phi(\mu/\sigma)$, where $\mu = (\beta_k^k/8m_k) - (\beta_1^k/8m_1) - \theta_1^k$ and σ is the standard deviation of ψ_i. Similarly, the expected number of workers belonging to group k in language group k's labor market is $n^k(1 - \Phi(\mu/\sigma))$. Notice the similarity to Lemma 2.1.

Lemma A.1. *Under Assumption A.1, the following statements hold. In the majority group's labor market, the expected profit of a firm that has an address y on the unit circle is*

$$E(\pi_1) = \frac{n^1}{2m_1}(\alpha^1 - \frac{\beta_1^1}{4m_1}) + \sum_{k=2}^{K} \frac{n^k \cdot \Phi(\mu/\sigma)}{2m_1}[\alpha^k - \rho s \frac{\phi(\mu/\sigma)}{\Phi(\mu/\sigma)} - \frac{\beta_1^k}{4m_1}] - C_1. \quad \text{(A.5)}$$

In the minority group's labor market, the firm's expected profit is

$$E(\pi_k) = \frac{n^k \cdot (1 - \Phi(\mu/\sigma))}{2m_k}[\alpha^k + \rho s \frac{\phi(\mu/\sigma)}{1 - \Phi(\mu/\sigma)} - \frac{\beta_k^k}{4m_k}] - C_k. \quad \text{(A.6)}$$

Note $\mu = (\beta_k^k/8m_k) - (\beta_1^k/8m_1) - \theta_1^k$.

Proof. The proof is analogous to that of Lemma 2.2 and is omitted. □

The result is analogous to that of Lemma 2.2. Since minority workers in the majority labor market belong to several language groups, the firm's revenue is composed of contributions from minority workers of $K-1$ groups. Also notice that equation (A.6) actually represents $K - 1$ equations.

The zero-profits conditions then characterize the equilibrium number of firms $m_l, l = 1, ..., L(K)$. The Jacobian J associated with the zero-profits conditions is

$$J = \begin{bmatrix} \frac{\partial E[\pi_1]}{\partial m_1} & \frac{\partial E[\pi_1]}{\partial m_2} & \frac{\partial E[\pi_1]}{\partial m_3} & \cdots & \frac{\partial E[\pi_1]}{\partial m_K} \\ \frac{\partial E[\pi_2]}{\partial m_1} & \frac{\partial E[\pi_2]}{\partial m_2} & 0 & \cdots & 0 \\ \frac{\partial E[\pi_3]}{\partial m_1} & 0 & \frac{\partial E[\pi_3]}{\partial m_3} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial E[\pi_K]}{\partial m_1} & 0 & 0 & \cdots & \frac{\partial E[\pi_K]}{\partial m_K} \end{bmatrix}.$$

Note that except on the diagonal, the first column, and the first row, all other elements are zeros.

Proposition A.1. *If (1) Assumption A.1 holds, (2) the model parameters are such that there is at least one equilibrium* $(m_1, ..., m_L)$ *such that* $m_l > 0$ *for all* $l = 1, ..., L$*, (3)* $\rho = 0$*, and (4) Evaluated at the equilibrium* $(m_1, ..., m_L)$*,* J *is negative definite, the following comparative statics hold at the equilibrium* $(m_1, ..., m_L)$ *for* $k, k' = 2, ..., K$ *and* $k \neq k'$*.*

1. *As (a)* n^1*,* α^1*,* C_k*, or* β_k^k *increases, or (b)* θ_1^k*,* C_1*,* β_1^1*, or* β_1^k *decreases,* m_1 *increases. The comparative statics of* m_1 *with respect to* n^k*,* α^k*, and* σ *are ambiguous.*

2. *As (a)* n^1*,* α^1*,* C_k*, or* β_k^k *increases, or (b)* θ_1^k*,* C_1*,* β_1^1*, or* β_1^k *decreases,* m_k *decreases. The comparative statics of* m_k *with respect to* n^k*,* α^k*, and* σ *are ambiguous.*

3. *As (a)* C_k *or* β_k^k *increases, or (b)* θ_1^k *or* β_1^k *decreases,* $m_{k'}$ *decreases. The comparative statics with respect to* n^k*,* α^k*, and* σ *are ambiguous.*

Proof. It is straightforward to verify the continuous differentiability of the zero-profits equations. It is also true that $|J| \neq 0$ at the equilibrium because of the assumption. Therefore, the implicit function theorem holds and the equation system defines a set of implicit functions for m_l where $l = 1, 2, \cdots, K$. Following similar steps in the proof of Proposition 2.1, I get

$$\begin{bmatrix} \frac{\partial E[\pi_1]}{\partial m_1} & \frac{\partial E[\pi_1]}{\partial m_2} & \cdots & \frac{\partial E[\pi_1]}{\partial m_K} \\ \frac{\partial E[\pi_2]}{\partial m_1} & \frac{\partial E[\pi_2]}{\partial m_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial E[\pi_K]}{\partial m_1} & 0 & \cdots & \frac{\partial E[\pi_K]}{\partial m_K} \end{bmatrix} \begin{bmatrix} \frac{\partial m_1}{\partial \phi} \\ \frac{\partial m_2}{\partial \phi} \\ \vdots \\ \frac{\partial m_K}{\partial \phi} \end{bmatrix} = \begin{bmatrix} -\frac{\partial E[\pi_1]}{\partial \phi} \\ -\frac{\partial E[\pi_2]}{\partial \phi} \\ \vdots \\ -\frac{\partial E[\pi_K]}{\partial \phi} \end{bmatrix},$$

where ϕ represents any exogenous variable in the model. Using Cramer's rule, the derivative of m_k with respect to ϕ can be written

$$\frac{\partial m_k}{\partial \phi} = \frac{|J_k|}{|J|}, \qquad k = 1, 2, ..., K \tag{A.7}$$

where J_k is the new matrix formed by substituting the right-hand side column vector to the kth column of the Jacobian J. Note also that $\frac{\partial E[\pi_1]}{\partial m_k} = n^k \frac{\partial \Phi}{\partial m_k} \frac{\alpha^k - (\beta_1^k/4m_1)}{2m_1} < 0$ and $\frac{\partial E[\pi_k]}{\partial m_1} = -n^k \frac{\partial \Phi}{\partial m_1} \frac{\alpha^k - (\beta_k^k/4m_k)}{2m_k} < 0$ for all $k = 2, ..., K$.

I pick the comparative static of m_1, m_k, and $m_{k'}$ with respect to β_1^k as an example, where $k' \neq k$ and $k, k' = 2, ..., K$. It is easy to verify that $-\frac{\partial E[\pi_1]}{\partial \beta_1^k} > 0$, $-\frac{\partial E[\pi_k]}{\partial \beta_1^k} < 0$, and $-\frac{\partial E[\pi_{k'}]}{\partial \beta_1^k} = 0$.

First, I derive the result for m_1. I can calculate $|J_1|$ as

$$|J_1| = (-\frac{\partial E[\pi_1]}{\partial \beta_1^k} \frac{\partial E[\pi_k]}{\partial m_k} - \frac{-\partial E[\pi_k]}{\partial \beta_1^k} \frac{\partial E[\pi_1]}{\partial m_k}) \prod_{l \neq 1,k} \frac{\partial E[\pi_l]}{\partial m_l}.$$

Using assumptions that $\frac{\partial E[\pi_k]}{\partial m_k} < 0$ and the results that $-\frac{\partial E[\pi_1]}{\partial \beta_1^k} > 0$, $-\frac{\partial E[\pi_k]}{\partial \beta_1^k} < 0$, and $\frac{\partial E[\pi_1]}{\partial m_k} < 0$, we get that $sign(|J_1|) = sign((-1)^{K-1})$. We know that $sign(|J|) = sign((-1)^K)$, so $\frac{\partial m_1}{\partial \beta_1^k} < 0$.

Next, I derive the result for $m_{k'}$, where $k' \neq k$. I calculate $|J_{k'}|$ as

$$|J_{k'}| = -\frac{\partial E[\pi_k]}{\partial m_1} (-\frac{\partial E[\pi_1]}{\partial \beta_1^k} \frac{\partial E[\pi_k]}{\partial m_k} - \frac{-\partial E[\pi_k]}{\partial \beta_1^k} \frac{\partial E[\pi_1]}{\partial m_k}) \prod_{l \neq 1,k',k} \frac{\partial E[\pi_l]}{\partial m_l}.$$

Using similar argument as in the first case, we get $sign(|J_{k'}|) = sign((-1)^{K-2})$, so $\frac{\partial m_{k'}}{\partial \beta_1^k} > 0$.

Lastly, I derive the result for m_k. This case is more complicated. The determinant of J_k is

$$\begin{aligned} |J_k| &= -\frac{\partial E[\pi_k]}{\partial \beta_1^k} (\prod_{l \neq k} \frac{\partial E[\pi_l]}{\partial m_l} - \sum_{l \neq 1,k} [\frac{\partial E[\pi_l]}{\partial m_1} \frac{\partial E[\pi_1]}{\partial m_l} \prod_{j \neq 1,k,l} \frac{\partial E[\pi_j]}{\partial m_j}]) \\ &- \frac{-\partial E[\pi_1]}{\partial \beta_1^k} \frac{\partial E[\pi_k]}{\partial m_1} \prod_{l \neq 1,k} \frac{\partial E[\pi_l]}{\partial m_l}. \end{aligned}$$

We can verify that $\prod_{l \neq k} \frac{\partial E[\pi_l]}{\partial m_l} - \sum_{l \neq 1,k} [\frac{\partial E[\pi_l]}{\partial m_1} \frac{\partial E[\pi_1]}{\partial m_l} \prod_{j \neq 1,k,l} \frac{\partial E[\pi_j]}{\partial m_j}]$ is the determinant of a submatrix of J. This submatrix is obtained by removing the kth row and kth column of J. Because J is negative definite, the sign of the determinant of this submatrix is $(-1)^{K-1}$. It is the feedback effect from other minority markets to minority market k. Using previous assumptions and results, we get that $sign(|J_k|) = sign((-1)^K)$. As a result, $\frac{\partial m_k}{\partial \beta_1^k} > 0$. Though tedious, it is straightforward to get all the comparative statics contained in the proposition. $\square$

Some results are familiar. For example, as majority population increases, the number of firms in the majority labor market also increases, and the number of firms in any minority labor market decreases.

However, some results require more delicate thought. Take the comparative statics with respect to the language learning cost θ_1^k as an example. It is clear that the number of majority firms increases as θ_1^k decreases. Because more workers from language group k enter the majority labor market, the majority firm can earn a positive profit. More firms will enter the majority labor market.

This change will in turn affect other minority labor markets, too. Because there are more firms in the majority labor market, the wages there are also higher. Workers from language group $k' \neq k$, for example, will leave their own language group's labor market. As a result, the number of firms in labor market k' decreases. The change in language group k is then transmitted to labor market k'. I call it a "ripple effect".

How does the number of firms in labor market k change? This case is even more complicated. We know that the numbers of firms in all other minority labor markets decrease. These changes have feedback effect to labor market k. It is reasonable to believe these changes increase the wages in the majority labor market further. Even smaller number of firms will operate in labor market k. This is shown in the proposition.

However, it is possible the feedback from other minority labor markets acts to decrease the wages in the majority labor market in some extreme cases. Therefore, workers from language group k leave the majority labor market. Larger number of firms operate in labor market k. This complicated feedback effect is reflected in the complicated expression of the determinant of J_k in the proof. Obviously, feedback effect is another form of the so-called "ripple effect".

A.3 Definitions of Variables

Language at Home: There are two types of home languages reported in the 2001 Canadian Census: the language "most often" spoken and the language "spoken on a regular basis". I use a combination of the two variables to represent a person's home language. For example, if a person reports speaking English "most often" but reports speaking non-official languages "on a regular basis", I deem her as a non-official language speaker and assign her to the appropriate language group. Since the Census classify those who report multiple non-official languages into the category "other". I employ the knowledge of non-official language indicators to find those individuals' home languages.

Language at Work: Its construction is similar to that of language at home.

Education: It is calculated based on the "highest level of schooling" reported in the data. I augment this variable by "years of university", "highest grade of elementary or secondary schooling", "highest degree, certificate, or diploma", and "years of college education" to get a more precise measure of schooling. More details are available upon request.

Experience: It is calculated as the person's education plus 7. If the person finishes at most 9 years of schooling, his experience is his age minus 16.

Sex, Marital Status: I get these variables from the data directly.

Age at Immigration Dummy: It equals 1 if a person reports that he immigrated when he was older than 19. It equals 0 if he reports otherwise, or if he reports that he was born in Canada. For non-immigrants who were not born in Canada, I use their actual ages to construct the dummy variable.

Occupation Dummies: I use the National Occupation Classification developed by Statistics Canada and Human Resources Development Canada to identify the occupations.

Instrument #1 for minority language group population: The first instrument is the log distance (in km's) between the place from which the individual's language originates and the CMA where the individual lives. The place of origin is normally a capital city. I calculate the distance from the city to Vancouver and Toronto, respectively. I then calculate the distance from the individual's CMA to Vancouver and Toronto. I choose the shorter route from the place of origin to the individual's CMA among two routes, the route via Vancouver and the route via Toronto. This distance approximates the flight distance between the place of language origin to the individual's CMA.

Instrument #2 for minority language group population: The second instrument is a predicted population measure based on the work by Card (1999). I first calculate the percentage accounted for by a CMA of the total national population of a language group in 1990.[32] I multiply this ratio by the language group's country-wide population in 2000. The result is a predicted population of the language group in a CMA.

Instrument for majority language group population: The construction of this instrument is similar to that of instrument #2 for minority language group population.

[32]Census 1991 is the earliest census that has the same language categories as those in Census 2001.

A.4 Two Types of Wage Gaps for First Generation Only

Table A.1 reports the estimated Within-Labor-Market Wage Gap for first generation immigrants only. It does not include visible minority status in the regressions. Based on Model 3 for all CMAs, the wage gap is 4.5% but not statistically significant. However, it is higher compared with Model 2, which does not include the Inverse Mill's ratio.

This result is consistent with the ability sorting explanation. Those who enter the majority market tend to have higher ability, so they are likely to have higher wages. Therefore, the raw estimate is an underestimation of the wage gap. The selection correction should increase the estimated wage gap as a result. The coefficient before the selection correction is positive, which is also consistent with ability sorting.

Table A.2 reports the same specifications as those in Table 2.5 for first generation immigrants. Model 3 for all CMAs shows a 12.8% Within-Labor-Market Wage Gap, which is higher than the estimated gap of 12.4% based on Model 2. Again, this result is consistent with the ability sorting explanation.

Table A.3 reports the same specifications of Table 2.6 for first generation immigrants. The estimated Within-Language-Group Wage Gap based on Model 3 is higher than that based on Model 2. This is still inconsistent with the ability sorting explanation. The ability sorting explanation predicts that the selection correction should reduce the estimate. We should note that the estimated wage gap of 26% is not significantly different from zero.

Table A.4 reports the same specifications of Table 2.7 for first generation immigrants. The estimated Within-Language-Group Wage Gap based on Model 3 is -2.7%, lower than that based on Model 2. This is consistent with the ability sorting explanation. However, the estimate based on Model 2 is not statistically different from zero.

Table A.1: Within-Labor-Market Wage Gap for 1st Generation Ignoring Discrimination

Specification	All CMAs			CMAs outside Quebec		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Majority group D_i	0.1286^a	0.0364	0.0451	0.1257^a	0.0177	0.0182
	(0.0271)	(0.0695)	(0.0702)	(0.0273)	(0.0713)	(0.0719)
Education	0.0623^a	0.0374^a	0.0399^a	0.0594^a	0.0355^a	0.0358^a
	(0.0046)	(0.0034)	(0.0041)	(0.0052)	(0.0037)	(0.0044)
Experience	0.0145^a	0.0176^a	0.0172^a	0.0143^a	0.0172^a	0.0171^a
	(0.0024)	(0.0023)	(0.0024)	(0.0025)	(0.0024)	(0.0025)
Experience sqd.	-0.0001	-0.0001^a	-0.0001^a	-0.0001	-0.0001^b	-0.0001^b
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
(De-meaned Educ)D_i	0.0147^a	0.0071^c	0.0048	0.0167^a	0.0092^b	0.0089^c
	(0.0051)	(0.0037)	(0.0042)	(0.0057)	(0.0042)	(0.0046)
(De-meaned Exp)D_i	0.0132^a	0.0081^b	0.0087^a	0.0135^a	0.0083^b	0.0084^b
	(0.0033)	(0.0031)	(0.0031)	(0.0035)	(0.0034)	(0.0033)
(De-meaned Exp^2)D_i	-0.0002^a	-0.0001^c	-0.0001^c	-0.0003^a	-0.0001^c	-0.0001^b
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Sex (Female=1)	-0.1389^a	-0.1235^a	-0.1236^a	-0.1413^a	-0.1201^a	-0.1201^a
	(0.0138)	(0.0132)	(0.0133)	(0.0146)	(0.0141)	(0.0141)
Married	0.1329^a	0.1208^a	0.1204^a	0.1359^a	0.1258^a	0.1257^a
	(0.0221)	(0.0186)	(0.0188)	(0.0237)	(0.0200)	(0.0201)
Sex*Married	-0.0471^b	-0.0478^a	-0.0476^a	-0.0472^b	-0.0517^a	-0.0516^a
	(0.0190)	(0.0158)	(0.0159)	(0.0206)	(0.0176)	(0.0177)
Language score		0.0253	0.0472		0.0405	0.0436
		(0.0635)	(0.0702)		(0.0680)	(0.0750)
Immigrate after 19		-0.1242^a	-0.1281^a		-0.1233^a	-0.1238^a
		(0.0091)	(0.0102)		(0.0095)	(0.0105)
Inverse Mill's ratio			0.0802			0.0100
			(0.1081)			(0.1087)
Occ. dummies	No	Yes	Yes	No	Yes	Yes
N	44482	43678	43678	40598	39912	39912
R^2	0.0759	0.1111	0.1112	0.0740	0.1092	0.1092

The dependent variables are log hourly wages. The sample includes all individuals who were born outside Canada. The inverse Mill's ratio is computed using model 2 of Table 2.3 for all CMAs and CMAs outside Quebec, respectively. ^{a} significance level of 1%. ^{b} significance level of 5%. ^{c} significance level of 10%. Standard errors are in parentheses. The regression error terms are clustered by individual language group within a CMA.

Table A.2: Within-Labor-Market Wage Gap for 1st Generation Considering Discrimination

Specification	All CMAs			CMAs outside Quebec		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Majority group D_i	0.1307^a	0.1239^c	0.1275^c	0.0676	0.0500	0.0425
	(0.0407)	(0.0684)	(0.0669)	(0.0418)	(0.0683)	(0.0709)
Visible minority V_i	-0.1191^a	-0.1081^a	-0.1081^a	-0.1396^a	-0.1157^a	-0.1154^a
	(0.0379)	(0.0253)	(0.0253)	(0.0412)	(0.0265)	(0.0266)
Interaction D_iV_i	-0.0570	-0.0063	-0.0058	0.0302	0.0651^b	0.0636^c
	(0.0437)	(0.0315)	(0.0321)	(0.0476)	(0.0327)	(0.0331)
Education	0.0530^a	0.0290^a	0.0300^a	0.0596^a	0.0368^a	0.0341^a
	(0.0046)	(0.0044)	(0.0044)	(0.0044)	(0.0041)	(0.0051)
Experience	0.0170^a	0.0169^a	0.0167^a	0.0222^a	0.0210^a	0.0215^a
	(0.0031)	(0.0032)	(0.0032)	(0.0036)	(0.0032)	(0.0034)
Experience sqd.	-0.0001	-0.0001^c	-0.0001^c	-0.0002^a	-0.0002^a	-0.0002^a
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
(De-meaned Educ)D_i	0.0143^a	0.0087^b	0.0077^c	0.0224^a	0.0137^a	0.0164^a
	(0.0040)	(0.0034)	(0.0040)	(0.0043)	(0.0037)	(0.0052)
(De-meaned Exp)D_i	0.0108^a	0.0076^b	0.0079^b	0.0177^a	0.0130^a	0.0124^a
	(0.0034)	(0.0033)	(0.0036)	(0.0043)	(0.0038)	(0.0040)
(De-meaned Exp^2)D_i	-0.0002^a	-0.0001^b	-0.0002^b	-0.0004^a	-0.0003^a	-0.0003^a
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
(De-meaned Educ)V_i	0.0170^a	0.0136^a	0.0137^a	0.0034	0.0041	0.0039
	(0.0052)	(0.0047)	(0.0047)	(0.0054)	(0.0050)	(0.0050)
(De-meaned Exp)V_i	0.0012	0.0031	0.0031	-0.0092^b	-0.0055	-0.0054
	(0.0030)	(0.0032)	(0.0032)	(0.0044)	(0.0040)	(0.0040)
(De-meaned Exp^2)V_i	-0.0001^c	-0.0001^c	-0.0001^c	0.0001	0.0001	0.0000
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Language score		-0.0656	-0.0562		-0.0507	-0.0773
		(0.0628)	(0.0709)		(0.0673)	(0.0723)
Immigrate after 19		-0.0974^a	-0.0991^a		-0.0968^a	-0.0931^a
		(0.0111)	(0.0137)		(0.0138)	(0.0145)
Inverse Mill's ratio			0.0346			-0.0950
			(0.1228)			(0.1283)
Occ. Dummies	No	Yes	Yes	No	Yes	Yes
N	44482	43678	43678	142812	142089	142089
R^2	0.0824	0.1140	0.1140	0.1225	0.1468	0.1468

The dependent variables are log hourly wages. The sample includes all individuals who were born outside Canada. Sex, Marital Status, and an interaction of the two are also included as regressors. The inverse Mill's ratio is computed using model 2 of Table 2.3 for all CMAs and CMAs outside Quebec respectively. ^{a} significance level of 1%. ^{b} significance level of 5%. ^{c} significance level of 10%. Standard errors are in parentheses. The regression error terms are clustered by individual language group within a CMA.

Table A.3: Within-Language-Group Wage Gap for 1st Generation Ignoring Discrimination

Specification	All CMAs			CMAs outside Quebec		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Majority market L_i	0.1917^a	0.1578^a	0.2602	0.1953^a	0.1580^a	0.2458
	(0.0274)	(0.0244)	(0.1951)	(0.0301)	(0.0274)	(0.2008)
Education	0.0371^a	0.0178^a	0.0115	0.0348^a	0.0166^a	0.0112
	(0.0077)	(0.0055)	(0.0079)	(0.0087)	(0.0059)	(0.0084)
Experience	0.0130^a	0.0166^a	0.0205^a	0.0131^a	0.0165^a	0.0198^a
	(0.0040)	(0.0037)	(0.0044)	(0.0043)	(0.0039)	(0.0046)
Experience Sqd.	-0.0001	-0.0002^b	-0.0003^a	-0.0001	-0.0002^b	-0.0003^a
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
(De-meaned Educ)L_i	0.0260^a	0.0194^a	0.0333^a	0.0255^a	0.0188^a	0.0304^a
	(0.0054)	(0.0047)	(0.0057)	(0.0058)	(0.0051)	(0.0059)
(De-meaned Exp)L_i	0.0033	0.0031	-0.0023	0.0032	0.0030	-0.0017
	(0.0033)	(0.0032)	(0.0040)	(0.0035)	(0.0033)	(0.0041)
(De-meaned Exp^2)L_i	0.0000	0.0000	0.0001	0.0000	0.0000	0.0001
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Sex (Female=1)	-0.1364^a	-0.1237^a	-0.1229^a	-0.1417^a	-0.1196^a	-0.1193^a
	(0.0321)	(0.0305)	(0.0302)	(0.0350)	(0.0320)	(0.0317)
Married	0.0824^a	0.0842^a	0.0837^a	0.0813^a	0.0870^a	0.0865^a
	(0.0218)	(0.0197)	(0.0197)	(0.0232)	(0.0211)	(0.0211)
Sex*Married	-0.0196	-0.0221	-0.0220	-0.0165	-0.0241	-0.0237
	(0.0331)	(0.0273)	(0.0267)	(0.0355)	(0.0286)	(0.0280)
Language Score		0.0965^c	0.1166^b		0.1071^c	0.1260^b
		(0.0537)	(0.0489)		(0.0563)	(0.0516)
Immigrate after 19		-0.1225^a	-0.1379^a		-0.1263^a	-0.1391^a
		(0.0163)	(0.0271)		(0.0171)	(0.0288)
Correction for $L_i = 1$			0.2604			0.2077
			(0.1747)			(0.1765)
Correction for $L_i = 0$			0.1700			0.1412
			(0.1079)			(0.1108)
Occ. dummies	No	Yes	Yes	No	Yes	Yes
N	31223	30629	30629	28088	27587	27587
R^2	0.0536	0.0884	0.0891	0.0523	0.0873	0.0878

The dependent variables are log hourly wages. The sample includes all individuals who were born outside Canada. The correction terms are computed using model 2 of Table 2.3 for all CMAs and CMAs outside Quebec respectively. ^{a} significance level of 1%. ^{b} significance level of 5%. ^{c} significance level of 10%. Standard errors are in parentheses. The regression error terms are clustered by individual language group within a CMA.

Table A.4: Within-Language-Group Wage Gap for 1st Generation Considering Discrimination

Specification	All CMAs			CMAs outside Quebec		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Majority Market L_i	0.1061^a	0.0806^b	-0.0273	0.0894^b	0.0667^c	-0.1246
	(0.0358)	(0.0356)	(0.2141)	(0.0355)	(0.0362)	(0.2124)
Visible Minority V_i	-0.2181^a	-0.2129^a	-0.1809^a	-0.2441^a	-0.2352^a	-0.2191^a
	(0.0513)	(0.0536)	(0.0645)	(0.0504)	(0.0549)	(0.0645)
Interaction L_iV_i	0.0942^b	0.1069^b	0.0706	0.1144^b	0.1244^b	0.1074^c
	(0.0444)	(0.0487)	(0.0602)	(0.0458)	(0.0511)	(0.0604)
Education	0.0211^a	0.0030	0.0045	0.0165^a	-0.0004	0.0038
	(0.0068)	(0.0060)	(0.0077)	(0.0062)	(0.0056)	(0.0077)
Experience	0.0112^b	0.0124^b	0.0138^b	0.0112^c	0.0125^b	0.0127^c
	(0.0057)	(0.0057)	(0.0066)	(0.0060)	(0.0061)	(0.0071)
Experience Sqd.	-0.0001	-0.0001	-0.0001	-0.0001	-0.0001	-0.0001
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
(De-meaned Educ)L_i	0.0258^a	0.0191^a	0.0261^a	0.0264^a	0.0196^a	0.0229^a
	(0.0040)	(0.0039)	(0.0048)	(0.0041)	(0.0042)	(0.0048)
(De-meaned Exp)L_i	0.0051	0.0052	0.0030	0.0051	0.0053	0.0045
	(0.0034)	(0.0033)	(0.0045)	(0.0035)	(0.0034)	(0.0045)
(De-meaned Exp^2)L_i	-0.0000	-0.0000	0.0000	-0.0000	-0.0000	-0.0000
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
(De-meaned Educ)V_i	0.0281^a	0.0242^a	0.0242^a	0.0302^a	0.0266^a	0.0266^a
	(0.0063)	(0.0055)	(0.0055)	(0.0062)	(0.0055)	(0.0055)
(De-meaned Exp)V_i	0.0058	0.0070	0.0065	0.0058	0.0064	0.0059
	(0.0049)	(0.0052)	(0.0053)	(0.0051)	(0.0056)	(0.0058)
(De-meaned Exp^2)V_i	-0.0002^b	-0.0002^b	-0.0002^b	-0.0002^c	-0.0002^c	-0.0002^c
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Correction for $L_i = 1$			0.2641			0.2277
			(0.1657)			(0.1651)
Correction for $L_i = 0$			-0.0146			-0.0813
			(0.1118)			(0.1082)
Occ. dummies	No	Yes	Yes	No	Yes	Yes
N	31223	30629	30629	28088	27587	27587
R^2	0.0627	0.0936	0.0938	0.0622	0.0929	0.0930

The dependent variables are log hourly wages. The sample includes all individuals who were born outside Canada. The correction terms are computed using model 2 of Table 2.3 for all CMAs and CMAs outside Quebec respectively. Though not reported in this table, Sex, Marital Status, their interaction, language score, and immigration age dummy are also included as regressors. ^{a} significance level of 1%. ^{b} significance level of 5%. ^{c} significance level of 10%. Standard errors are in parentheses. The regression error terms are clustered by individual language group within a CMA.

Appendix B

Appendix for Chapter 3

B.1 Housing Market Equilibrium

Within a neighborhood, the price of housing must adjust to equal supply and demand, given the distribution of population across neighborhoods $\{n_j^k\}_{k,j}$. The Marshallian demand for housing by an individual of group k in neighborhood j is denoted $h(r_j, y^k, q_j^k)$. Remember that y^k is a function of r, the vector of neighborhood housing prices including the price of the composite good, in the exchange economy. Also recall that q_j^k is a function of $(n_j^1, ..., n_j^K)$. Therefore, the housing demand can be reexpressed as $h_j^k(r)$, given a vector of populations $(n_j^1, ..., n_j^K)$. The market clearing condition in neighborhood j is

$$\sum_{k=1}^{K} n_j^k h_j^k(r) = S_j, \text{where } j = 1, 2, ..., J. \tag{B.1}$$

The system of equations in equation (B.1) can then be reexpressed as

$$\sum_{k=1}^{K} n_j^k h_j^k(r) - \sum_{k=1}^{K} n_j^k w_j^k = 0, \text{where } j = 1, 2, ..., J, \tag{B.2}$$

where I substitute the endowments w's into the equation. In addition, the market demand for the composite good must equal the total endowment of the composite good in the city.

$$\sum_{k=1}^{K} \sum_{j=1}^{J} n_j^k z_j^k(r) - \sum_{k=1}^{K} L^k w_0^k = 0. \tag{B.3}$$

The above equations (B.2) and (B.3) can be reexpressed in matrix form as

$$e(r) = 0, \tag{B.4}$$

where $e(r)$ is the excess aggregate demand vector for the composite good and the housing in neighborhood $j = 1, ..., J$.

Proposition B.1. *If the utility function $u(z, h, q_j^k)$ is continuous, strictly quasiconcave, and strictly increasing in h, a vector of strictly positive price vector $r \in R_{++}^J$ exists to clear all the neighborhood housing markets conditional on a population distribution $\mathbf{N} \equiv (n_1^1, ..., n_1^K, ..., n_j^k, ..., n_J^1, ..., n_J^K)$,*

Proof. The proof follows closely the standard proof for existence of Walrasian equilibrium. See Proposition 17.B.2 and Proposition 17.C.1 in Chapter 17 of the book by Mas-Colell, Whinston, and Green (1995). Here, $e(r)$ corresponds to the excess demand function in their discussions.

The steps of the proof is as follows. First, I need to establish that $e(r)$ is continuous and homogeneous of degree zero. I also need to establish that $r \cdot e(r) = 0$, or r satisfies the Walras' Law. There are two other more technical properties. These are listed in Proposition 17.B.2 of Mas-Colell, Whinston, and Green (1995).

The continuity and homogeneity property are inherited directly from the properties of the Marshallian Demand for housing $h(r)$. $h(r)$ is continuous because the utility function is continuous, strictly monotone, and strictly quasiconcave. $h(r)$ is homogeneous of degree zero because multiplying the price vector by a constant does not change the choice set faced by individuals. The Walras's Law comes directly from the strict monotonicity of the preference. See Chapter 3 of Mas-Colell, Whinston, and Green (1995) for details.

Given these conditions analogous to those listed in Proposition 17.B.2 of Mas-Colell, Whinston, and Green (1995), Proposition 17.C.1 implies the existence of r that clears all the housing market. The proof of Proposition 17.C.1 involves the construction of fixed-point correspondence and the use of Kakutani's Fixed Point Theorem. Because $e(r)$ is homogeneous of degree zero in r, we can focus on price vectors in the unit simplex in R^J, which is nonempty, compact, and convex. These properties of the unit simplex allows us to use the fixed-point theorem. □

Proposition B.1 establishes that there is a vector of housing prices r given a population distribution vector $\mathbf{N}$. However, there may be multiple r's for one $\mathbf{N}$. The following proposition establishes the uniqueness of equilibrium housing price vector r. The uniqueness of r facilitate establishing the existence of inter-neighborhood equilibrium.

Proposition B.2. *If the vector of excess aggregate demand $e(r)$ is continuously differentiable and its Jacobian matrix has strictly positive off-diagonal terms and strictly negative diagonal terms, $e(r) = 0$ has a unique solution.*

Proof. It is equivalent to Proposition 17.F.3 of Mas-Colell, Whinston, and Green (1995), so the proof is omitted. Also see Definition 17.F.2 in Mas-Colell, Whinston, and Green (1995) and the discussions that follows for the conditions used to derive the result. □

Proposition B.2 implies that we can express the neighborhood housing price r_j as a function of $\mathbf{N}$. We can then reexpress ν^{kj} as a function of the population distribution vector $\mathbf{N}$.

B.2 Inter-Neighborhood Equilibrium

An inter-neighborhood equilibrium attains if the proportion of people from group k who actually choose neighborhood j or n_j^k/L^k equals the probability prescribed by the utility maximization model $P^{kj}(\nu^{k1}(\mathbf{N}), ..., \nu^{kJ}(\mathbf{N}))$. It is immediately clear that this represents a fixed-point problem. More formally, the equilibrium is a distribution of the population $\{n_j^k \geq 0\}_{i,k}$ such that the following conditions hold

$$\sum_{i=1}^{J} n_j^k = L^k > 0 \text{ for all } k = 1, ..., K, \tag{B.5}$$

$$n_j^k = P^{kj}(\nu^{k1}(\mathbf{N}), ..., \nu^{kJ}(\mathbf{N}))L^k \text{ for all } k = 1, ..., K \text{ and } i = 1, ..., I, \tag{B.6}$$

where $\mathbf{N} = (n_1^1, ..., n_1^K, ..., n_j^k, ..., n_J^1, ..., n_J^K)$.

Proposition B.3. *If $\nu^{kj}(\mathbf{N})$ is continuous in n_j^k for $0 \leq n_j^k \leq L \equiv \sum_{k=1}^{K} L^k$, and $P^{kj}(\nu^{k1}, ..., \nu^{kJ})$ is continuous in ν^{kj} for all $k = 1, ..., K$ and for all $j = 1, ..., J$, an inter-neighborhood equilibrium exists.*

Proof. The proof follows closely the proof of Theorem 1 in Miyao (1978b). Each n_j^k is an element of the following simplex

$$n_j^k \geq 0 \text{ and } \sum_{i=1}^{J}\sum_{k=1}^{K} n_j^k = L > 0. \tag{B.7}$$

It is well known that a simplex is nonempty, compact, and convex, and hence satisfies the requirement for applying Brouwer's fixed point theorem. The function $P^{kj}L^k$ maps from the above simplex to itself. It remains to confirm the continuity of the mapping P^{kj}. Since continuity is preserved under function composition, the conditions in the proposition implies that P^{kj} is continuous in $\mathbf{N}$. □

The proposition establishes the existence of an inter-neighborhood equilibrium. However, it tells us little about whether the equilibrium is unique or stable. Miyao (1978b) discussed the conditions under which such claims can be made.

B.3 Descriptions of Data

B.3.1 Description of Data and Variables for Section 3.2

I use Canadian Ethnic Diversity Survey (2003) to construct a dataset to analyze the determinants of a person's social networks' ethnic composition. Below, Table B.1 summarizes the basic characteristics of the variables I constructed and used to do the analysis. The sample is restricted to people whose mother tongue is not one of the official languages and who are of first or second generation.

Table B.1: Variables from Ethnic Diversity Survey(2003)

Variable	Mean	Std. Dev.	N
Diversity of Ethnicity of a Person's Friends	2.972	1.2	10048
Dummy - Speak Official Languages with Parents till 15	0.142	0.349	10898
Dummy - Speak Official Languages with Siblings till 15	0.372	0.483	10639
Education	13.122	3.037	10511
Age	44.192	17.861	10961
Age Squared	2271.944	1672.763	10961
Generation Status	1.303	0.46	10961
Own Group's Ratio of City Population	0.062	0.057	10961
Dummy - Immigrate before Age 5	0.059	0.237	10841
Dummy - Immigrate between Age 6-14	0.11	0.313	10841
Dummy - Immigrate between Age 15-24	0.194	0.395	10841
Dummy - Immigrate after Age 44	0.28	0.449	10841

B.3.2 Description of Data and Variables for Section 3.4

Table B.2 list the summary statistics for the housing attributes I use in the regressions. It comes from the housing transaction data provided by British Columbia Assessment Authority.

Table B.3 summarizes the basic statistics for various socio-economic variables used in my analysis. It is constructed from Census Profile Tables for Vancouver in Census Years of 1991, 1996, and 2001.

Table B.2: Summary Statistics of the Housing Transaction Data

Variable	Mean	Std. Dev.	N
House Price	280919.44	202541.684	57851
Number of Bedrooms	3.896	1.246	57677
Number of Full Bathrooms	1.645	0.959	57851
Number of Half Bathrooms	1.004	0.869	57851
Age of the House	20.825	20.786	57698
Age Squared	865.719	1541.52	57698
House has a Pool	0.034	0.182	57851
Number of Parking Space	1.454	0.796	57851
House has a Deck	0.674	0.469	57851
House has a Basement Suite	0.174	0.379	57851
House in a Waterfront Lot	0.003	0.052	57851
Prime View	0.023	0.149	57851
Good View	0.012	0.108	57851
Fair View	0.011	0.105	57851

Table B.3: Summary Statistics of Census Tracts in 1991, 1996, and 2001

	2001		1996		1991	
Variable	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Average Housing Value	292135	123518	322031	142298	244629	102374
Herfindahl Index of All Groups	0.555	0.168	0.585	0.174	0.656	0.164
Herfindahl Index of Minority Groups	0.331	0.189	0.312	0.179	0.238	0.122
Number of Bedrooms	2.708	0.733	2.626	0.725	2.678	0.688
Average Age of the Housing Stock	28.113	10.606	26.814	10.441	24.933	10.633
Share of Apartments	0.391	0.3	0.374	0.294	0.335	0.297
Ownership Ratio	0.647	0.203	0.633	0.206	0.624	0.227
Share of New Immigrants	0.391	0.141	0.483	0.146	0.299	0.113
Share of University Graduates	0.219	0.116	0.176	0.105	0.143	0.095
Share of People above Age 65	0.122	0.06	0.12	0.059	0.122	0.066
Average Household Incomes	65567	22839	56839	18669	52952	16550
N	386		298		298	

Appendix C

Appendix for Chapter 4

C.1 Model Parametrization

Table C.1: The Parameter Setup

Parameter Name	Symbol	Initial Value	Range of Variation
Mean of Terminal Dividend	μ	1.191	n/a
Variance of Ter. Dividend	σ^2	1%	$[0.1\%, 1.9\%]$
Absolute Risk Aversion	a	4	$[0.1, 10]$
Variance of Noise Trading	σ_z^2	6.25%	$[0.156\%, 15.625\%]$
Precision of the Signal	σ_y^2/σ_z^2	0.5	$[0.01, 0.99]$
Initial Endowment of Asset	$\bar{z}$	4.7754	n/a
Fraction of Informed	λ	0.1	$[0.01, 1]$

C.2 Proofs and Derivation

C.2.1 Proof of Proposition 4.1

Let us see the informed trader's problem first. Rewriting the problem in expression 4.10 in terms of the fundamental random variables y, $\tilde{z}_{n1}$, $\tilde{z}_{n2}$, and $\tilde{d}$, the informed investor maximize

$$\max_{z_{i1}} \quad E[-e^{-a[\tilde{z}_{i2}(\tilde{d}-\mu+a\sigma^2\tilde{z}_{i2})+z_{i1}(\mu-a\sigma^2\tilde{z}_{i2}-p_1)+W_{i1}]}|y,p_1],$$
$$\text{where} \quad \tilde{z}_{i2} = \bar{z} - \tilde{z}_{n1} - \tilde{z}_{n2}. \tag{C.1}$$

Suppose that the informed trader expects p_1 to be a linear function of y and $\tilde{z}_{n1}$, which I will show the existence later. The informed investor can then deduce the value of the realized value of $\tilde{z}_{n1}$ from observing both y and p_1. Therefore, $\tilde{z}_{n1}$ is no longer random and $\tilde{z}_{i2} = \bar{z} - \tilde{z}_{n1} - z_{n2}$.

Evaluating the expectation in equation (C.1) involves quadratic function of normal random variables. It can be shown that for a multivariate normally distributed $n \times 1$ vector $\mathbf{x}$ with mean μ and covariance matrix Σ or precision matrix $\mathbf{H} = \Sigma^{-1}$, two constants a and f, one $n \times 1$ vector $\mathbf{v}$, and one

$n \times n$ symmetric matrix $\mathbf{Q}$, the following relation holds as long as $a\mathbf{Q} + \mathbf{H}$ is positive definite and symmetric.

$$E[\exp[-a(f + \mathbf{v}^T x + \frac{1}{2}\mathbf{x}^T\mathbf{Q}\mathbf{x})]] = |\mathbf{H}|^{1/2}|a\mathbf{Q} + \mathbf{H}|^{-1/2} \times \\ \exp[-(af + \frac{1}{2}\mu^T\mathbf{H}\mu - \frac{1}{2}(a\mathbf{v} - \mathbf{H}\mu)^T(a\mathbf{Q} + \mathbf{H})^{-1}(a\mathbf{v} - \mathbf{H}\mu))]. \qquad (C.2)$$

See Christensen and Feltham (2002) Appendix 3A for the detailed derivation.

Given the distributional assumptions in section 4.2.3, it is easy to show that $E([\tilde{d}-\mu,\ \tilde{z}_{i2}]^T|y, p_1) = [0,\ y]^T$, and $Var([\tilde{d}-\mu,\ \tilde{z}_{i2}]^T|y, p_1) = \begin{bmatrix} \sigma^2 & 0 \\ 0 & \sigma_\epsilon^2 \end{bmatrix}$. Using the above information and the formula provided before, we can derive the certainty equivalent of the terminal wealth $\tilde{W}_{i3}$ as

$$CE_{\tilde{W}_{i3}} = z_{i1}(\mu - p_1) + \frac{1}{2a}y^2\sigma_\epsilon^{-2} - \frac{a(z_{i1}\sigma^2 + m_{i1}a^{-2}\sigma_\epsilon^{-2})^2}{2(\sigma^2 + a^{-2}\sigma_\epsilon^{-2})}, \qquad (C.3)$$

where $m_{i1} \equiv E[\tilde{z}_{i2}|y, p_1] = \bar{z} - y - z_{n1}$. The first order condition then implies equation (4.12).

The uninformed trader solves a similar problem to problem (C.1) with $\tilde{z}_{i2}$ and z_{i1} being replaced by the corresponding variables. Another difference lies in the conditioning set. The uninformed only knows p_1, but she rationally expects the price to be a linear function of the private information y and the noise trader's demand $\tilde{z}_{n1}$ at $t = 1$. Denote $\psi = y + (\pi_z/\pi_y)z_{n1}$ be a linear transformation of p_1. It is evident that ψ is informationally equivalent to p_1. Unlike the informed trader, the uninformed cannot infer from the price of the realized $\tilde{z}_{n1}$ anymore because she cannot observe y in the first place.

The posterior mean of the random vector $[\tilde{d} - \mu,\ \tilde{z}_{u2}]^T$ given ψ is then $[0,\ \frac{\sigma_y^2 + \frac{\pi_z}{\pi_y}\sigma_z^2}{\sigma_\psi^2}\psi + \bar{z}]^T$, where $\sigma_\psi^2 = \sigma_y^2 + (\frac{\pi_z}{\pi_y})^2\sigma_z^2$. The posterior covariance matrix of the random vector $[\tilde{d} - \mu,\ \tilde{z}_{u2}]^T$ conditional on ψ is $\begin{bmatrix} \sigma^2 & 0 \\ 0 & 2\sigma_z^2 - \frac{\sigma_y^4 + (\frac{\pi_z}{\pi_y})^2\sigma_z^4}{\sigma_\psi^2} \end{bmatrix}$. Following similar steps as those for informed traders, I can derive the demand from an uninformed trader as that in equation (4.13).

The market clearing condition is

$$\lambda z_{i1} + (1 - \lambda)z_{u1} = \bar{z} - z_{n1}. \qquad (C.4)$$

Substituting the expressions of z_{i1}, z_{u1}, and p_1 in proposition 4.1 into the left hand side of the above equation, we can verify that the left hand side exactly equals the right hand side. Therefore, we've found a rational expectations equilibrium price. We've proved the existence of a linear pricing equilibrium by constructing such an equilibrium price.

The steps to follow in constructing this equilibrium price is as follows. First, substitute the expressions for z_{i1} and z_{u1} into the market clearing condition. Second, collect terms and solve p_1 as a function of y, z_{n1}, and other parameters. Lastly, equalize the coefficient before y to π_y and the coefficient before z_{n1} to π_z to get two equations in π_y and π_z. In the end, we get the exact form of p_1 in proposition 4.1.

C.2.2 Proof of Corollary 4.1 and Proposition 4.2

Proof of Corollary 4.1:

Proof. If no one is informed at $t = 1$, the demand from a typical rader is then

$$z_{u1} = \frac{\mu - p_1}{a\sigma^2} + \frac{h_z}{a^2\sigma^2}\frac{1}{a\sigma^2}\Big(\mu - a\sigma^2(\bar{z} - z_{n1}) - p_1\Big), \tag{C.5}$$

where $h_z = 1/\sigma_z^2$ since no one is informed. Also note that z_{n1} is no longer a random variable for the trader because she deduce it from the observed price, knowing the exact form of the pricing function. Substitute this into the market clearing condition $z_{u1} = \bar{z} - z_{n1}$, we can get $p_1|_{\lambda=0} = \mu - a\sigma^2(\bar{z} - z_{n1})$. □

Proof of Proposition 4.2:

Proof. Observing the maximization problem in expression (4.16), we note the analogy of this problem to the problem solved in proposition 4.1. Using the joint distribution assumptions on $\tilde{d}$, $\tilde{z}_{n1}$, and $\tilde{z}_{n2}$ and employing the formula to evaluate the expectation of the exponential-quadratic function, we can find the certainty equivalent of $\tilde{W}_{u3}$. The first order condition then implies the optimal choice of z_0. The market clearing condition implies the equilibrium price p_0. The exact forms of the two are shown in Proposition 4.2. □

C.2.3 Proof of Proposition 4.5

The trading volume at $t = 1$ when there is no private information can be expressed as $T_1|_{\lambda=0} = \frac{1}{2}(|\tilde{z}_{n1}| + |\tilde{z}_{u1} - \bar{z}|)|_{\lambda=0}$. Note the the first term in the parentheses is the number of sales per rational trader from the noise traders and the second term is the number of shares bought by a typical rational trader. One half the sum of the absolute values of those two numbers is then the trading volume. We know $\tilde{z}_{u1}|_{\lambda=0} = \bar{z} - \tilde{z}_{n1}$ if no one receives private information. Therefore, the trading volume $T_1|_{\lambda=0} = |\tilde{z}_{n1}|$. Since the expectation of the absolute value of a zero mean normal random variable is $(\frac{2}{\pi})^{\frac{1}{2}}$ times the standard deviation of the random variable, $E(T_1|_{\lambda=0}) = (\frac{2}{\pi})^{\frac{1}{2}}\sigma_z$. Similarly, we can derive $E(T_2|_{\lambda=0}) = (\frac{2}{\pi})^{\frac{1}{2}}\sigma_z$.

The trading volume at $t = 1$ and $t = 2$ can be derived respectively as $T_1|\lambda = \frac{1}{2}(|\tilde{z}_{n1}| + \lambda|\tilde{z}_{i1} - \bar{z}| + (1-\lambda)|\tilde{z}_{u1} - \bar{z}|)$ and $T_2|\lambda = \frac{1}{2}(|\tilde{z}_{n2}| + \lambda|\tilde{z}_{i2} - \tilde{z}_{i1}| + (1-\lambda)|\tilde{z}_{u2} - \tilde{z}_{u1}|)$. In the above, $\tilde{z}_{i2} = \tilde{z}_{u2} = \bar{z} - \tilde{z}_{n1} - \tilde{z}_{n2}$, $\tilde{z}_{i1} = \bar{z} + \phi_y\tilde{y} + \phi_z\tilde{z}_{n1}$, and $\tilde{z}_{u1} = \bar{z} + \delta_y\tilde{y} + \delta_z\tilde{z}_{n1}$, where ϕ_y, ϕ_z, δ_y, and δ_z are as expressed in Proposition 4.5 and derived from the results in Proposition 4.1. The two trading volumes can then be reexpressed as $T_1|\lambda = \frac{1}{2}(|\tilde{z}_{n1}| + \lambda|\phi_y\tilde{y} + \phi_z\tilde{z}_{n1}| + (1-\lambda)|\delta_y\tilde{y} + \delta_z\tilde{z}_{n1}|)$ and $T_2|\lambda = \frac{1}{2}(|\tilde{z}_{n2}| + \lambda|\tilde{z}_{n2} + (1+\phi_z)\tilde{z}_{n1} + \phi_y\tilde{y}| + (1-\lambda)|\tilde{z}_{n2} + (1+\delta_z)\tilde{z}_{n1} + \delta_y\tilde{y}|)$. Employing the above mentioned formula for calculating the expectation of the absolute value of a normal random variable and using the the joint distribution assumptions in Section 4.2.3, we can get the expressions for the expected trading volumes contained in Proposition 4.5.