Two Essays on Spatial Skill Sorting and Household Saving Behavior

by

Minjuan Sun

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved April 2018 by the
Graduate Supervisory Committee:

Todd Schoellman, Co-Chair
Gustavo Ventura, Co-Chair
Galina Vereshchagina

ARIZONA STATE UNIVERSITY

May 2018

ABSTRACT

This dissertation studies two wide ranging phenomena and their socio-economic impacts: urban divergence in terms of geographical skill sorting and fast rising housing prices. The first essay explores the empirical pattern as well as the driving forces behind the American cities' diverging path over the past forty years. Compared to the rest of the U.S. cities, the top 20 largest cities have been growing faster in several aspects, such as city-average wage, housing price, and measured innovation intensity (e.g., patents, venture capital). In addition, this geographical divergence has contributed substantially to the rising inequality in America. To explore the causes of this divergence, this paper constructs a spatial sorting model where entrepreneurs with different talents can freely move across cities. The key idea is that cities with advantages in innovation attract more productive entrepreneurs and more workers, thereby driving up wages and housing prices. Two things distinguish my models from others: 1. Large cities are having endogenous innovation advantage in equilibrium; 2. I can freely explore the driving forces behind the divergence, with an emphasis on how technology changes can reinforce the spatial sorting mechanism. Specifically, three types of technological changes have increased the benefits of skill clustering in innovative cities: general productivity increases; improvements in communications technologies; and declines in trade costs.

The second essay studies how heterogeneous households respond to the fast rising housing prices through their life-cycle behaviors. Chinese housing market has been undergoing a rapid booming period since 1998, causing the house prices increasing significantly. As a result, households endured severe financial burdens to buy homes at price-to-income ratios of around six. Along with the rising house prices, household savings rate has been increasing consistently since 1998. Can the rising house prices be an important factor to explain the increase in household

saving rate? This paper develops a life cycle dynastic model with endogenous choice on housing, coresidence and intergenerational transfer, then quantitatively analyze the effect of housing price on household saving. It shows that housing is an important motive for saving, and it accounts for about 35% of the increase in household savings rate. The housing situation affects households' saving behavior through three channels. First, households are financially constrained due to the down payment requirement and they choose to limit their consumption in order to buy houses. Second, young adults live in their parents' houses for a long time and save more intensively, since they get to pay less for the housing expenses under coresidence. Thirdly, older parents make large sum of intergeneration transfer in aid of the children's housing purchase, indicating the housing affordability issue also has influence on old parents' saving decisions.

TABLE OF CONTENTS

iv

LIST OF FIGURES

LIST OF TABLES

Chapter 1

Unequal Cities: Innovation, Skill Sorting and Inequality

1.1 Introduction

Over the past three decades, the spatial distribution of skills have become increasingly more dispersed, with the robust feature that the initially skilled cities are becoming more and more skilled[1]. The process of differently skilled individuals being sorted into different cities is called spatial sorting. The direct result of spatial sorting is the phenomenon some refer to as "the Great Divergence" among American cities. Successful cities have attracted more and more highly-talented people and enabled them to work collaboratively. Meanwhile, these cities are experiencing higher population, wage and housing price growth than the rest of cities. With the fast advancement in technology, people's capability for long-distance communication and shipping goods over long distances have vastly improved. Therefore the need for spatial concentration should have decreased. However, exactly the opposite is happening in the real world. Behind this increased trend of skill sorting and urban inequality, lies a puzzling question. Why do people, especially the most talented people, keep moving towards large cities, without regard to their extremely high living costs?

In modern society, the role of cities in accelerating the flow of technology and ideas has taken a central place. The geographic density provided by cities brings people together, and this proximity stimulates idea spread and enables collaboration. Talented people cluster not simply because they like each other's company

---

[1]Many papers have studied this phenomenon, such as Goldin and Katz (2007), Ganong and Shoag (2016), Moretti (2012) and Diamond (2016).

or they prefer urban centers with far superior amenities, but because they can enjoy productive advantages and knowledge spillovers that such concentrations bring. As of 2013, about 53% of the top 5% earners in U.S. metropolitan areas are located in the top 20 large cities. This is an increase from 39% in 1980[2]. Meanwhile, there is unprecedented concentration of innovation activities happening in these cities. About half of the new product innovations in the 1980s occurred in just four metropolitan areas: Boston, New York, San Francisco, and Los Angeles[3]. And almost all VC investments were made in major cities, see figure 1.1.

Figure 1.1: **Spatial concentration of US patenting and venture capital**
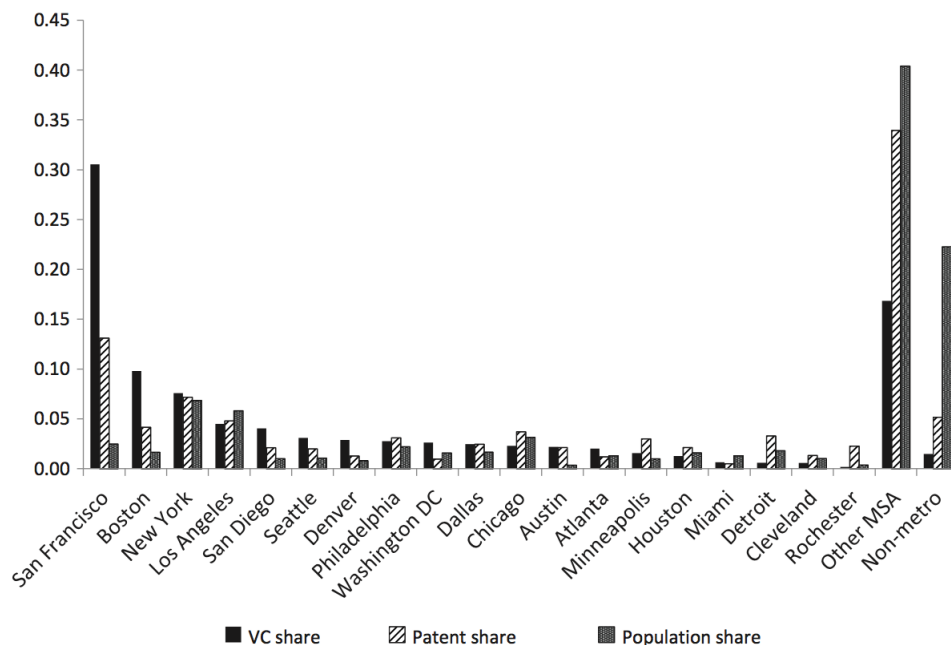
Note: VC calculations use the share of deals over the 1990-2005 period. Patent calculations use the share of granted patents applied for from each city during 1990–2005. The population share is from 1999. The figure is taken from Chatterji et al. (2014). Source data are from VentureXpert, USPTO patent data, and county level population statistics.

To explain the increasing trend of spatial skill sorting and innovation concen-

---

[2]The metropolitan population share in these 20 cities have increased from 22% in 1980 to 29% in 2013.

[3]This empirical findings about spatial concentration of the commercialization of innovation can be found in Feldman and Audretsch (1999). For R&D activity, Buzard and Carlino (2013) showed that the spatial concentration of establishments undertaking R&D is more pronounced than firms.

tration, this paper formulates a multi-city model of innovation and production. The model allows for two types of agents: production workers and entrepreneurs. Production workers are kept relatively simple: they simply produce and consume. Entrepreneurs are monopolistic competitors who differ in productivity[4]. Entrepreneurs run firms and innovate, which increases their productivity. If an entrepreneur decides to innovate, he pays a cost in return for a random productivity draw from the entrepreneurial productivity distribution of the aggregate economy.

Both workers and entrepreneurs choose which of two cities to live and work in. The cities are identical except that one (generically, the first city) has an endowed advantage in aiding innovation. I think of this policy as capturing rules, regulations, and the endowment of complementary inputs such as input suppliers, universities, and so on[5]. This endowed advantage of the first city is balanced by congestion costs that are in equilibrium stronger, inducing a non-trivial sorting problem. Every entrepreneur wants to live in the city with better innovation environment, since it offers them the best chance to improve their productivities. But only the more talented ones are able to survive the competition. The complementarity between agglomeration economies—innovation in this paper—and entrepreneurs' productivity leads to the sorting of more skilled entrepreneurs into larger cities. In equilibrium, they invest more resources on innovation and pay

---

[4]The setting of entrepreneurs are similar to that of the standard heterogenous firm model in Merlitz (2003).

[5]Traditional discussions of natural advantages focus on geographic features such as harbors and coal mines. The rise of New York in the early nineteenth century is the result of its central location and protected harbor which made it the natural hub of shipping and immigration system. But, the technology advancement has rendered these geographic features less relevant over time. Bleakley and Lin (2012) believed that the early developed cities have some long-lived assets that could coordinate contemporary investment and cause a persistent effect. Apart from these, factors such as local policies or proximity to universities and financial institutes can be important as well. For instance, some believe Silicon Valley and Boston became important centers for innovation in part because of their proximity to Stanford University and MIT (Lee and Nicholas, 2013). And the lack of "noncompete and nondisclosure clauses" policy do set California apart from other states. The "noncompete and nondisclosure clauses" restrict workers from starting new businesses that could possibly in direct competition with their old employers.

higher wages to workers. In summation, the interactions between innovation, sorting, and agglomeration economies shape the income distribution and exacerbate inequality across cities of different sizes.

I start by showing that the model equilibrium allows us to understand the facts outlined above. The large city attracts more talented entrepreneurs who innovate more, leading to a higher average productivity. Those more productive entrepreneurs pay higher wages, which then attract more workers to city 1. This establishes a link between productivity and city size. On the other hand, the agglomeration forces are inevitably met with congestion forces: higher housing prices and negative amenity caused by overcrowding. Furthermore, this model generates two straightforward implications on spatial inequality. First, the productivity premium of the large city, resulted from skill sorting and intense innovation, leads to a wage gap across cities. Second, the large city is more unequal comparing to the small one, because entrepreneurs in large city benefit the most from both the superior innovation environment and larger market share.

Next, I use the comparative statics of the model to establish some insights on possible driving forces behind the increased sorting by cities over the last fifty years. In particular, I show that a simple exogenous rise in total factor productivity (TFP) can induce stronger skill sorting. Whenever aggregate TFP rises, every city becomes more productive, indicating that every city's wage and housing price will rise as well. But the average wage and housing price in in large city increase more, thus leading to a tough selection. In response, the marginal entrepreneurs in large city will be better off moving to small city, which then raises both city's productivity. But the productivity in big city disproportionately increases more because of the complementarity between innovation and entrepreneurial productivity. In this paper, the aggregate TFP is described by the distribution of entrepreneurs' productivity. Since productivity follows a Pareto distribution in the

4

model, an increase in TFP can be proxied by an increase in the minimum productivity threshold, which effectively depicts the catching up of the least talented entrepreneurs in small cities (Tonetti and Perla, 2014). Due to the fat tail property of Pareto distribution, the productivity in large city actually increase more. Therefore, the equilibrium productivity and wage gaps across different sized cities grow over time.

In the baseline model, the result of perfect skill sorting hinges on the assumption that one city has exogenous innovation advantage. One way to remedy this is to include localized knowledge spillover effect. In section 1.4, I extend the model by assuming that cities have no fundamental differences, instead the process of innovation is slightly altered. Now knowledge spread is more likely to occur if agents live in the same place, since face-to-face meeting is more effective in idea exchange. The most talented entrepreneurs choose the same city, which in equilibrium becomes the large city. The relatively less talented entrepreneurs, on the other hand, cannot afford to live in the big cities due to its higher cost.

In addition, as information technology (internet, email, video chat) has changed the way people interact with each other, so it is interesting to analyze how these telecommunications affect the value of cities as well as the pattern of skill sorting. Depending on whether information technology and city are complement or substitute with each other[6], the effects vary. If they are complement, telecommunications improve big cities innovation environment to a greater degree, thus the value of big city increases. As a result, more and more relatively unskilled individuals will move towards big cities. On the other hand, if they are substitute, telecommunications improve small cities innovation environment to a greater

───────────────────────────

[6]According to Gaspar and Glaeser (1996), when telecommunications technology improves, there are two opposing effects on cities and face-to-face interactions: some relationships that used to be face-to-face will be done electronically (an intuitive substitution effect), and some individuals will choose to make more face-to-face contacts(an complementary effect). So far, there is no consensus on this subject.

degree, and the relatively unskilled individuals will move towards small cities instead.

Lastly, I study how inter-regional trade affects the composition of skill as well as income distribution across cities. I extend the model by including bilateral trade. The results suggest that skill sorting and inequality become greater with the decline of iceberg trade costs. Intuitively, entrepreneurs in big cities benefit more from trade due to the asymmetrical trade costs. A decrease in trade costs disproportionately increases the market size as well as wage in large cities. These changes are more beneficial to the more productive entrepreneurs than to the less productive ones. Therefore, the least talented entrepreneurs, originally living in large cities, will be better off moving to small cities instead. In equilibrium, spatial skill sorting becomes stronger in response to trade costs decrease. As a result, the productivity as well as wage gaps become wider.

Overall, the model proved to be tractable enough to study the multi-factor spatial equilibrium system. It offers some insights on what caused the geographic divergence among American cities with respect to productivity, wage and inequality, and what factors contribute to the deeper and growing trend in spatial skill sorting. In addition, I also stress the role of TFP increase and trade in increasing the relevant market size, which reinforces skill sorting, sustains larger and more productive cities.

1.1.1 Related Literature

There are three strands of literature closely related to this paper. The first relevant literature is about the cross-sectional skill sorting across cities (Behrens and Robert-Nicoud (2014); Behrens, Duranton, and Robert-Nicoud (2014); Eeckhout, Pinheiro, and Schmidheiny (2014); Gennaioli, La Porta, Lopez-de Silanes,

6

and Shleifer (2013)). Behrens, Duranton, and Robert-Nicoud (2014) developed a multi-city model that explained the complementarities between skill sorting, occupation selection, and agglomeration economies. They successfully replicated the stylized facts about sorting, agglomeration, and selection in cities. The main force of agglomeration in their model comes from scale effect: the entrepreneurial productivity gain increases with city population size. This approach proves to be effective in analyzing city's cross sectional skill sorting and size distribution. Plus the selection of entrepreneurs within each city provides plenty of productivity overlapping across cities, which is a nice feature, because in real world skill sorting is imperfect. This paper is built on their framework, and it has one major difference: the productive advantage (or innovation advantage) of large cities is allow to be endogenous.

The second relevant literature is focuses on localized knowledge spillover effects on individuals' spatial choices (Henderson (1974); Duranton and Puga (2004)). Knowledge spillover is directly related to the acquisition of skills and the learning of new technologies, hence it is one of the most important mechanisms giving rise to agglomeration economies. Most recently, Davis and Dingel (2016) developed a spatial equilibrium framework to show why skill premiums are higher in large cities. Similar to this paper, they believed that localized idea exchange is the main agglomeration force and large cities have better idea-exchange environments. Compared to their paper, this paper not only can answer the question why large cities are capable of attracting more skilled labor force cross-sectionally, but also can answer the question what are the driving forces behind the great divergence across cities over the past three decades.

Lastly, there are a number of papers discussing innovation, economic growth and inequality (Lucas and Moll (2014); Perla and Tonetti (2014); Perla, Tonetti, and Waugh (2016); Gabaix et al. (2016) and Jones and Kim (2017)). Lucas and Moll

Table 1.1: **Skill Composition Changes between 1990 and 2010**

	Share of College (Including Above)			Share of Above College	
	Small City	Big City		Small City	Big City
1990	19.82%	24.99%	2000	7.57%	8.72%
2010	25.91%	32.78%	2010	8.21%	10.91%
Growth	6.09%	7.79%	Growth	0.64%	2.19%

(Note: Small city refers to MSAs with less or equal to 1 million population, whereas big city refers to the ones with more than 1 million population.)

(2014) explored a model of human capital and ideas exchange that gives rise to endogenous growth. Aghion et al. (2017) showed that innovation and top income inequality are positively correlated within U.S. states and across U.S. commuting zones. Jones and Kim (2017) developed a model in which entrepreneurs expend effort to increase their productivities while new entrants replace incumbents in a process of creative destruction generates a Pareto distribution for top incomes. It suggests economic forces that raise top income inequality: forces that increase the return of fast-growing entrepreneurs' innovation effort. Their findings can be reconciled with this paper. Empirically, top earners exhibit a geographic pattern, with the richest disproportionally locating in large cities. This paper suggests that the complementarity between top entrepreneurs' innovation effort and locational fundamentals lead to the spatial pattern of innovation concentration as well as top income inequality.

1.2 Motivating Evidence

Empirical evidence that motivating this paper is that big cities have higher productive advantages for firms, and higher income premium for workers. It is even better if there are evidence to suggest that the main benefits of living in big cities come from knowledge spillover effect.

la Roca and Puga (2017) was an important empirical study of the positive relationship between city size and labor income as well as income growth. They studied Spain's labor market by cities, and found that workers in big cities not only have higher earnings, they also have higher growth rates in earnings. The interpretation of income growth is human capital accumulation through experience. Their result indicates that the experience accumulated in bigger cities is substantially more valuable than experience accumulated in smaller cities, and these experience is even more valuable for workers with higher ability. And it directly prove that there are important learning benefits to working in bigger cities which is embedded in workers' human capital. The implication of their analysis is an important premise for this paper: the advantages of living in big cities come from the opportunities they provide for individuals to learn from others, and the learning benefit is greater for people with higher abilities.

Figure 1.2: **Share of Top 5% Earners Located in Big Cities**
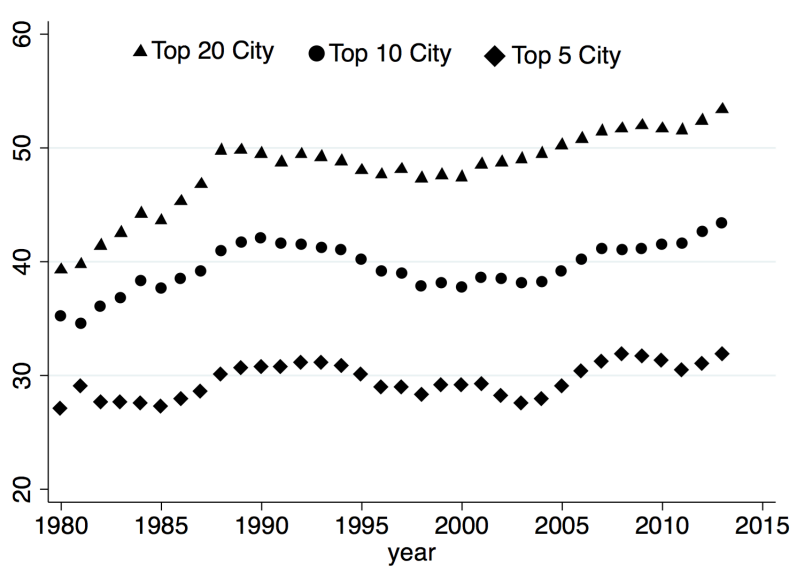
Note: Calculations are based on US Census Bureau data for 366 MSAs between 1980 and 2013. The figure plots the share of top 5% earners locating in subgrouped cities out of MSAs.

There are many ways that make cities different: size, productivity and living cost. But most fundamentally, cities differ in the composition of human capital.

9

Figure 1.3: **Income Inequality and Wage Gap**

Fig. a: Income Ratio between top 5% and bottom 25%   Figure b: City Average Hourly Wage (Dollar)
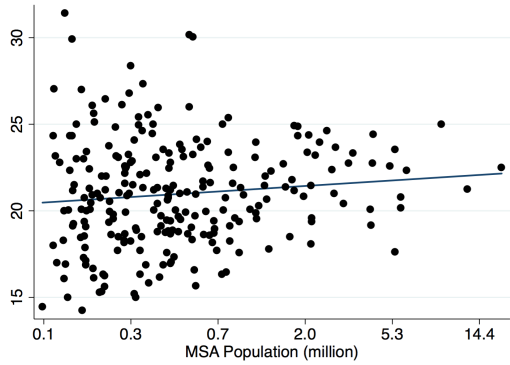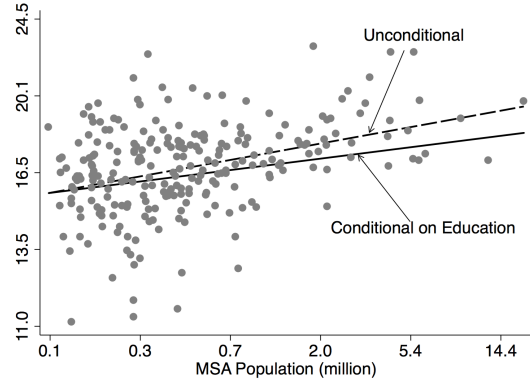
In recent literature, skill sorting has been a robust feature from the data. According to Combes et al. (2008), human capital composition differences across cities can explain up to 40–50% of the size-productivity relationship. Table 1.1 shows that the share of skilled labor force in big cities is disproportionately higher. Additionally, the increase of skilled labor share in big cities between 1990 and 2010 is greater than that of small cities, implying that skill sorting indeed becomes stronger over time. Another way to examine the pattern of skill sorting is to check the share of top talents located in different sized city. If income is a good indicator for individual skill/talent, figure 2.1 shows that the share of most talented population (proxied by top 5% earners in all the MSAs) in big cities are increasing over year. The increasing share of top earners in big cities might come from two channels: rich people moving into big cities or the income growth of top earners in big cities is greater than that of the rest of places. This paper intends to explore what forces have caused the upward trend in skill sorting, and analyze the direct consequences of increasing sorting on urban inequality.

City size is not only correlated with labor force skill composition, it is also correlated with inequality in economic outcomes. The fact that large cities are more unequal is a robust feature of the data, and it has been discussed in papers,

like Glaeser et al. (2009) and Baum-Snow and Pavan (2013). There are two aspects of urban inequality concerned in this paper, which are shown in figure 1.3 : wage inequality across cities (right panel) and top income inequality within cities (left panel). Skill sorting, interacting with agglomeration force, has significant effect on wage inequality across different sized cities. Meanwhile, the force that induced skill sorting, innovation or knowledge spillover, plays a more direct role in top income inequality. The most talented entrepreneurs who are sorted into big cities benefit more from innovative activities, thus allowing them disproportionately the top earnings.

1.3 The Simple Model

This section describes and solves the multi-city theoretical framework. The premise of this model is similar to Behrens et al. (2014) and Desmet and Rossi-Hansberg (2014). There are two types of agents: worker and entrepreneur. Agents live for one period, and they are free to move across cities. In particular, they have preferences over good consumption and housing. I will describe the production and agent decisions in turn.

1.3.1 City

This paper considers an economy with two cities $j \in \{1, 2\}$. There is a fixed housing supply, $\bar{H}_j$, which is owned by landlords who only consume final consumption goods. Aside from housing supply, cities are different with respect to only one locational characteristic: the first city (generically) has an endowed advantage in aiding innovation. The innovation advantage of city 1 is going to play an

important role in what type of agents are going to be sorted into each city.

1.3.2 Agent's Problem

As introduced above, the economy includes two types of agents: workers and entrepreneurs. There are L identical workers, who consume and produce. And there is a continuum of entrepreneurs of mass Ω, who differ in productivity. Each entrepreneur is a monopolist who produces a differentiated intermediate good. Based on productivity, entrepreneurs choose where to locate. After moving into a city, they can also innovate, which allows them to produce their variety more productively.

1.3.2.1 Worker's Problem

All workers are endowed with one unit of labor, which they supply inelastically. In each city, workers' total income is equal to the nominal wage, which is spent on final consumption good (serving as numeraire) and housing. The utility function for a worker facing wage w_j and housing price P_{hj} in city j is Cobb-Douglas form [7]:

$$U^W(j) = \max_{C_j, H_j} \left(\frac{C_j}{\mu}\right)^{\mu} \left(\frac{H_j}{1-\mu}\right)^{1-\mu} - a_j;$$

$$s.t.\ C_j + P_{hj}H_j = w_j. \tag{1.3.1}$$

---

[7]For empirical evidence using U.S. data in support of the constant housing expenditure share implied by the Cobb-Douglas functional form, see Davis and Ortalo-Magne (2011).

where C_j represents final goods consumption, H_j represents housing consumption, and a_j is city-specific congestion amenity. In this paper, congestion amenity refers to the amenities resulting from overcrowding[8], such as long commuting time, pollution, crime, or simply the difficulty to find a parking space. Because agents are perfectly mobile across cities, and all workers are ex-ante identical by assumption, then workers will be indifferent between living in two cities.

1.3.2.2 Entrepreneur's Problem

The entrepreneur's preference is similar to that of worker. The major difference is that entrepreneurs earn profits by providing intermediate variety, whereas workers earn nominal wages. Because innovation is a random process, so entrepreneur's profit is uncertain. But, I can still define the expected utility of entrepreneur i, who chooses to live in city j and has an initial productivity z:

$$U_i^e(j,z) = \max_{C_{ji},H_{ji}} \left(\frac{C_{ji}}{\mu}\right)^\mu \left(\frac{H_{ji}}{1-\mu}\right)^{1-\mu} - a_j; \qquad (1.3.2)$$

$$s.t.\ C_{ji} + P_{hj}H_{ji} = \Pi_{ji}\left(z\right).$$

Where $\Pi_{ji}\left(z\right)$ stands for expected net profit, which is the result of endogenous choices on location, innovation and production decisions. In the next section,

---

[8]This model assumes that congestion amenity is fixed, rather than an increasing function of population size. If I adopt a different approach, such that $a_j = L_j^\rho$, with ρ representing the congestion elasticity w.r.t city population. I can still solve the model, except it is more difficult. The literature on the estimates of congestion elasticity ρ is quite sparse, so far, I am only aware of Combes et al. (2016). In Duranton et al. (2015), they believed that this congestion elasticity should be very small, which is close to 2%. However, the congestion cost in their definition mostly includes land and housing prices. In this paper, the congestion force resulted from housing price is included in variable P_{hj}. Therefore, the congestion elasticity in this paper should be even smaller, which indicates that as city's population size increases, its congestion amenity increases but only very barely.

the setup of entrepreneur's production and innovation problems are introduced in detail.

1.3.3 Production

The structure of production within a city is a two-step process: intermediate goods and final output good. In each city there is a final good producer that supplies the final output good competitively. The final good is produced by aggregating the mass of intermediate varieties that are provided by monopolistically competitive entrepreneurs. And it serves as numeraire in both cities, thus its price is set to 1. The final good producer chooses the quantity to purchase of each variety:

$$\max_{y_{ji}} \left(\int_{i \in \Omega_j} y_{ji}^{\theta} di \right)^{\frac{1}{\theta}}, \ 0 < \theta < 1 \tag{1.3.3}$$

$$s.t. \int_{i \in \Omega_j} p_{ji} y_{ji} di = Y_j.$$

where $\sigma = \frac{1}{1-\theta}$ is the elasticity of substitution; y_{ji} is the amount of intermediate good i used for final good production in city j, p_{ji} is the price of intermediate good i, and Y_j is the aggregate output of the final consumption good. The measure Ω_j defines the endogenous mass of entrepreneurs who choose to live and produce in city j. Entrepreneurs have claims on the net profits from selling varieties. The more productive entrepreneurs are more efficient in engaging production activity, thus having higher profits. (Thinking about the order of these things)

The interpretation of entrepreneur is not quite so literal. They are, in general, talented entrepreneurial types. They can be Silicon Valley startups, software

engineers, successful authors, or doctors with new techniques. Every entrepreneur faces three choices: where to live, how much to innovate, and how much to produce. The model is solved in a backward fashion, thus both production and innovation decisions are operated on the premise that entrepreneurs already choosing the optimal location. In intermediate goods production, labor is the only input. The output for intermediate good i in city j is given by:

$$y_{ji} = (q_{ji})^\gamma l_{ji}, \tag{1.3.4}$$

where l_{ji} is the labor demanded to produce intermediate good i, and q_{ji} is entrepreneur's realized productivity after innovation. I assume for the rest of the paper: $\gamma = \frac{1-\theta}{\theta} = \frac{1}{\sigma-1}$, with $\sigma \geq 1$. Note that this is just a simplifying assumption which makes profits a linear function of productivity q_{ji}. It can be relaxed with a bit more algebra. Because all the labor resource is used to produce intermediate goods, then the labor market constraint is :

$$L_1 + L_2 = L; \ L_j = \int_{i \in \Omega_j} l_{ji} di. \tag{1.3.5}$$

It means that out of the aggregate population L, there are L_1 workers choosing city 1 and L_2 workers choosing city 2 in equilibrium. Next I introduce the process of innovation. Entrepreneurs need to decide whether and how much to invest in innovation.

1.3.4 Innovation

Entrepreneurs have differentiated productivity, which can be understood as their skill level or talent. Each entrepreneur is identified by a draw of productivity z

from a Pareto distribution $G(\cdot)$, which is described by its cumulative distribution function (cdf) $\Pr(z > z_{min}) = 1 - \left(\frac{z_{min}}{z}\right)^{\alpha}$; where z_{min} is the minimum productivity threshold, and α is the Pareto tail parameter governing dispersion.

Upon choosing a location, each entrepreneur can improve his productivity through buying a chance of a new draw from distribution $G(\cdot)$. This innovation process can be understood as people learning from one another, which is similar to the interpretation in Lucas and Moll (2014)[9]. On the other hand, it can also be thought of as tangible or intangible investments that manifest themselves as improvements in productivity such as improved production practices, work practices, management practices, etc. see discussions in Holmes and Schmitz (2010).

Similar to the set-up of Desmet and Rossi-Hansberg (2014), an entrepreneur i with initial productivity z in city j can decide to buy a probability $\phi_{ji} \leq 1$ of innovating at cost $\psi(\phi_{ji} \mid z)$, which is paid out of the profit from selling variety. This process dictates that entrepreneur obtains an innovation with probability ϕ_{ji}, and with probability $(1 - \phi_{ji})$ his productivity is not affected by the investment in innovation. The entrepreneur who obtains a chance to innovate draws a new skill z^+ from distribution $G(z)$. The new productivity is adopted if it is higher than the initial productivity, $z^+ > z$; if not, entrepreneur will operate at his initial productivity level z. Then the expected productivity conditional on the initial z is:

$$E\left(z^+ \mid z, \ innovation\right) = \int_{z_{min}}^{\infty} \max\{x, z\} \, dG(x) = \frac{\alpha z}{\alpha - 1};$$

The added "plus" superscript refers to the productivity after the innovation decision. The expected productivity for entrepreneur i (with initial productivity

---

[9]In the baseline model, an entrepreneur in city 1 can improve his productivity by learning from entrepreneurs in city 2, without occurring extra innovation cost than he learns from those locating in the same city. An extended model of localized knowledge spillover is introduced in section 4, in which learning from distant cities is more costly than learning within the same city.

z) is:

$$q_{ji}(z) = \underbrace{\phi_{ji} \int_z^\infty x dG(x)}_{expected\ productivity\ if\ innovate} + \underbrace{(1 - \phi_{ji}) z}_{innovation\ not\ occur} \qquad (1.3.6)$$

Finally, I make some assumptions on the primitives of the model. I assume that the innovation cost function $\varphi(\phi_{ji} \mid z)$:

$$\frac{\partial \varphi(\phi_{ji} \mid z)}{\partial \phi_{ji}} > 0 \ and \ \frac{\partial \varphi^2(\phi_{ji} \mid z)}{\partial \phi_{ji}^2} > 0 \ for \ \phi_{ji} \in (0, 1).$$

and for any $z \geq z_{min}$:

$$\varphi(\phi_{1i} \mid z) < \varphi(\phi_{2i} \mid z). \qquad (1.3.7)$$

These two assumptions make sure that for any given productivity z: (i) the innovation cost is a convex function of innovation opportunity ϕ_{ji}, so that there is no corner solution problem; (ii) the innovation cost required in city 1 is lower than that in city 2, which is basically stating that the first city has endowed innovation advantage.

Before the model is solved, I need to formally define the spatial equilibrium. Given initial productivity distribution $G(z)$, city-specific innovation advantage χ_j and housing supply $\bar{H}_j$, a spatial equilibrium is a set of real functions $\{P_{hj}, w_j, C_j, H_j, l_{ji}, \phi_{ji}, Y_j, Q_j, L_j, \Omega_j\}$ of city j and entrepreneur i, such that:

- Given city-specific wage w_j and house price P_{hj}, workers choose consumption bundle $\{C_j, H_j\}$, and then locate optimally by solving problem (1.3.1);

- Ex-ante identical workers are indifferent between two cities, and the indifference condition leads to city's labor supply L_j, which is expressed in (1.4.12);

17

- Entrepreneurs choose optimal location by solving problem (1.3.2), and the mass of entrepreneurs choosing city j is characterized by Ω_j;

- Given initial productivity z and location choice, each entrepreneur chooses innovation opportunity ϕ_{ji} by solving problem (1.4.1);

- Given city's wage w_j, each entrepreneur chooses the number of workers l_{ji} to maximize profit;

- City's productivity Q_j is the average productivity of the mass of entrepreneurs Ω_j choosing city j, its expression is given in (1.4.10);

- The model assumes that the aggregate housing value is a constant share (λ) of the aggregate output Y_j[10]. Given housing price P_{hj}, the housing market clearing condition is $\bar{H}_j P_{hj} = \lambda Y_j$;

- Labor market clear, so L_j satisfied condition (1.3.5).

1.4 Model Results

Because an entrepreneur's decisions are rather complicated, I summarize them into four steps, which then serve as a roadmap for the actual solution process. Since I solve the model in a backward fashion, the process is as following:

1. Given city's wage w_j, entrepreneurs choose the optimal labor demand l_{ji} to maximize the expected profit $E\left(\pi_{ji} \mid z\right)$ from selling varieties. The detailed steps for solving the expected profit $E\left(\pi_{ji} \mid z\right)$ is given in appendix A1.

2. Second, given initial productivity z and location j, entrepreneur i chooses

---

[10]The assumptions on land market is similar to Davis and Nieuwerburgh (2014) and Redding and Rossi-Hansberg (016b).

innovation opportunity ϕ_{ji} to maximize the expected net profit $\Pi_{ji}(z)$:

$$\Pi_{ji}(z) = \max_{\phi_{ji}} E\left(\pi_{ji} \mid z\right) - \varphi\left(\phi_{ji} \mid z\right); \tag{1.4.1}$$

3. Third, with the expected net profit $\Pi_{ji}(z)$, entrepreneur i solves the consumption maximization problem described in (1.3.2), and get optimized expected utility $U_i^e(j, z)$.

4. Lastly, to decide the optimal location, he just needs to choose the city with higher utility:

$$\max_{j \in \{1,2\}} U_i^e(j, z). \tag{1.4.2}$$

For an entrepreneur with initial productivity z, if $U_i^e(1, z) > U_i^e(2, z)$, he chooses city 1; if $U_i^e(1, z) < U_i^e(2, z)$, he chooses city 2; if $U_i^e(1, z) = U_i^e(2, z)$, he is indifferent between living in two cities;

1.4.1 Optimal Innovation Choice

Having defined the equilibrium behavior of agents and the structure of the model. I can now describe the equilibrium properties. First, the city-specific wage, aggregate output, and profits are given by the following proposition.

Proposition 1: (Output, Wages, and Profits): Let w_j denote the wage rate in city j, and let π_{ji} denote the profit earned by entrepreneur i before innovation cost is paid. The equilibrium with monopolistic competition leads to

$$w_j = \theta Q_j^\gamma \tag{1.4.3}$$

$$Y_j = Q_j^\gamma L_j. \tag{1.4.4}$$

$$E\left(\pi_{ji} \mid z\right) = (1-\theta)Y_j \frac{q_{ji}(z)}{Q_j}. \tag{1.4.5}$$

where $Q_j = \int_{i \in \Omega_j} z_{ji}^+ di$ is the mean productivity across entrepreneurs locating in city j.

See proof in appendix A1. According to proposition 1, entrepreneur's individual profit from selling variety is proportional to local aggregate output, Y_j, and relative productivity, $\frac{q_{ji}(z)}{Q_j}$. Here Q_j represents city's average productivity, and its definition is: $Q_j = \int_{\Omega_j} q_{ji}(z)\, dz$. As introduced in section 1.3.4, innovation improves entrepreneur's productivity, and it happens with a chance ϕ_{ji}, which is endogenously determined by weighing the benefit and the cost of innovation. According to function (1.3.6), the expected productivity after innovation is:

$$q_{ji}(z) = \phi_{ji}\frac{\alpha z}{\alpha - 1} + (1 - \phi_{ji}) z \tag{1.4.6}$$

$$= \underbrace{\left(\frac{\phi_{ji} + \alpha - 1}{\alpha - 1}\right)}_{scale\ of\ innovation} z.$$

The above expression indicates that the benefit of innovation comes from the elevation of initial productivity z. And the scale of elevation, $\frac{\phi_{ji}+\alpha-1}{\alpha-1}$, solely depends on innovation opportunity ϕ_{ji} and Pareto tail parameter α. Now, it is time to solve entrepreneur's optimal innovation opportunity. To do that, I need to give the specific functional form for innovation cost:

$$\varphi\left(\phi_{ji} \mid z\right) = \frac{\chi_j w_j L_j}{1 - \phi_{ji}} \frac{z}{Q_j}; \tag{1.4.7}$$

where χ_j represents city-specific innovation efficiency parameter: smaller χ_j implies better innovation environment. To be consistent with the assumption that city 1 has a better innovation environment compared to city 2, I let $\chi_1 < \chi_2$[11]. Note that the innovation cost function is also increasing in city-specific wage (w_j), population size (L_i) and relative productivity $\frac{z}{Q_j}$. The increasing relationship between innovation cost and aggregate labor income $w_j L_j$ is to capture the labor cost in innovation[12]. Meanwhile, the more productive entrepreneurs are closer to technology frontier, thus have little room to improve, which is why their innovation cost is higher.

As we can see, innovation opportunity ϕ_{ji} is crucial in determining entrepreneur's location choice, since it governs the scale of productivity improvement from innovation. To determine the optimal innovation opportunity, entrepreneur i maximizes his expected net profit $\Pi_{ji}(z)$. The solution of innovation decision gives rise to the following expression:

$$\phi_{ji} = 1 - \left(\frac{\chi_j\left(\alpha - 1\right)}{\gamma}\right)^{1/2} = \phi_j^* \tag{1.4.8}$$

Because every entrepreneur in the same city has the same innovation opportunity, hence the expected productivity for entrepreneur i in city j is:

$$q_{ji}(z) = \underbrace{\left[\frac{\alpha}{\alpha - 1} - \left(\frac{\chi_j}{\gamma\left(\alpha - 1\right)}\right)^{1/2}\right]}_{scale\ of\ improvement} z. \tag{1.4.9}$$

---

[11]The rationale for one of the city having better innovation environment can be found in footnote 6.

[12]In equilibrium, the aggregate labor income $w_j L_j$ is proportional to the aggregate output, i.e. $Y_j = \frac{w_j L_j}{\theta}$. Then the defined innovation cost function indicates that innovation investment is proportional to entrepreneurial profit, since they are both proportional to local aggregate output.

This indicates that the productivity improvement from innovation is proportional, and the scale of improvement in city 1 is higher than city 2. Using the above expression, I can get city-specific average productivity:

$$Q_j = \int_{\Omega_j} \left(\frac{\phi_j^* + \alpha - 1}{\alpha - 1} \right) z_{ji} di = \left[\frac{\alpha}{\alpha - 1} - \left(\frac{\chi_j}{\gamma(\alpha - 1)} \right)^{1/2} \right] \bar{z}_j. \qquad (1.4.10)$$

where $\bar{z}_j = \int_{i \in \Omega_j} z_{ji} di$ is the mean of initial productivity from the mass of entrepreneurs Ω_j choosing to locate in city j. If the the mass of entrepreneurs choosing city j are relatively more productive, city j is more productive. Meanwhile, city's average productivity increases with the innovation efficiency parameter $1/\chi_j$, which is quite intuitive. Since the better the innovation environment (lower χ_j) is in a city, the more incentive entrepreneurs have to invest in innovation, as a result the scale of productivity improvement will be larger.

1.4.2 Equilibrium Spatial Sorting

This model involves two types of agents: worker and entrepreneur. Naturally, spatial sorting refers to the process both types deciding their optimal locations. I will start with worker's location choice first, since it is fairly easy.

1.4.2.1 The Sorting of Workers

Given city-specific housing price and wage, worker solves utility maximization problem defined in problem (1.3.1). With constant expenditure shares on final good and housing, worker's indirect utility function becomes:

$$U^W(j) = \frac{w_j}{P_{hj}^{1-\mu}} - a_j.$$

The first part of the indirect utility function $\frac{w_j}{p_{hj}^{1-\mu}}$ represents the real income after adjusting housing cost, the second part a_j represents the congestion amenity. Clearly, worker's location choice is determined by city's nominal wage and living cost, which includes housing price and local amenity. This simple utility form is able to accommodate a number of regional differences. The two cities differ in their relative productivity, with more productive city having higher nominal wage and higher housing price. In equilibrium, ex-ante identical workers are indifferent between living in two cities. Hence, worker's indirect utility function $U^W(j)$ equalizes:

$$\frac{w_1}{p_{1j}^{1-\mu}} - a_1 = \frac{w_2}{p_{h2}^{1-\mu}} - a_2 = u^*, \tag{1.4.11}$$

where u^* is the equilibrium common utility level. This indifference condition clearly states the tradeoff for choosing between two cities: the higher nominal wage in large city must be balanced out by a higher living cost. Using the housing market clearance condition $\bar{H}_j P_{hj} = \lambda Y_j$, house price can be expressed as a function of wage w_j and local population L_j:

$$P_{hj} = \frac{\lambda Y_j}{\bar{H}_j} = \frac{\lambda w_j L_j}{\theta \bar{H}_j}.$$

Let $\eta_j = \frac{\lambda}{\theta \bar{H}_j}$, then parameter $1/\eta_j$ can be perceived as land supply coefficient: city's fixed housing supply $\bar{H}_j$ is more restricted when η_j is higher. Replacing P_{hj} with $\eta_j w_j L_j$ in the indifference equation (1.4.11), the equilibrium labor supply in city j is

$$L_j = \frac{w_j^{\frac{\mu}{1-\mu}}}{\eta_j (u^* + a_j)^{\frac{1}{1-\mu}}}. \tag{1.4.12}$$

Equation (1.4.12) represents the labor supply in city j, which is increasing in city's wage rate, decreasing in land supply coefficient η_j and congestion amenity a_j.

1.4.2.2 The Sorting of Entrepreneurs

An entrepreneur's sorting problem is similar to that of the worker's. Each entrepreneur decides where to live based on three factors: expected entrepreneurial income, housing price and city-specific congestion amenity. With the optimal innovation opportunity expressed in (4.8), the equilibrium expected net profit is:

$$\Pi_{ji}(z) = E\left(\pi_{ji} \mid z\right) - \varphi\left(\phi_j^* \mid z\right) = \Gamma_j z, \qquad (1.4.13)$$

where $\Gamma_j = (1-\theta)\left[\frac{\alpha}{(\alpha-1)} - 2\left(\frac{\chi_j}{\gamma(\alpha-1)}\right)^{1/2}\right]\frac{Y_j}{Q_j}$ summarizes the benefit of an entrepreneur choosing city j, which is increasing in city's aggregate output Y_j and innovation efficiency parameter $1/\chi_j$. With entrepreneur's expected net profit as well as city's housing price, I can easily write out the optimal indirect utility function:

$$U_i^e(j, z) = \frac{\Gamma_j}{P_{hj}^{1-\mu}} z - a_j = \Psi_j\left(Q_j\right) z - a_j; \qquad (1.4.14)$$

in which

$$\Psi_j\left(Q_j\right) z = \frac{1}{\eta_j}\left(\frac{\theta}{u^* + a_j}\right)^{\frac{\mu}{1-\mu}}\left[\frac{\gamma\alpha}{(\alpha-1)} - 2\left(\frac{\gamma\chi_j}{(\alpha-1)}\right)^{1/2}\right] Q_j^{\frac{\gamma\mu+\mu-1}{1-\mu}} z$$

representing entrepreneur's expected real income after adjusting housing price. We can see there is a linear relationship between the expected real income and initial productivity z, indicating that the real income is higher in a more productive city. The reason for this is as following. On the one hand, a city is more productive if the mass of entrepreneurs in that city are on average more productive. On the other hand, entrepreneurs in city 1 benefit more from innovation, because they have higher innovation opportunity due to the endowed advantage. Moreover, the most talented entrepreneurs benefit the most from living in city 1. Therefore, there

24

exists a complementarity between innovation and productivity. In equilibrium, this complementarity is the driving force for agglomeration economy under this framework.

Due to the linearity between entrepreneur's indirect utility function $U_i^e(j, z)$ and his productivity z, there exists an unique skill threshold $\bar{z} > 0$, such that entrepreneur with initial productivity higher than $\bar{z}$ choosing to live in the more innovative city (city 1), and entrepreneurs with initial productivity lower than $\bar{z}$ choosing to live in city 2[13]. When city-specific amenities are unequal to each other, namely $a_1 > a_2$, the solution of above equation leads to the threshold for skill sorting:

$$\bar{z} = \frac{a_1 - a_2}{\Psi_1(Q_1) - \Psi_2(Q_2)}. \tag{1.4.15}$$

If this skill sorting threshold is larger than the minimum level z_{min} of the initial productivity distribution $G(\cdot)$, $\bar{z} > z_{min}$[14], there exists an unique spatial equilibrium characterized by perfect skill sorting. In addition, this equilibrium sorting result is efficient, because the most skilled entrepreneurs end up living in the city with innovation advantage that promote economic development. The following proposition states the existence and uniqueness of the spatial equilibrium. See proof in appendix A2.

Proposition 2: (Equilibrium Existence and Uniqueness): Assume $a_1 > a_2$, $\chi_1 < \chi_2$ and $\gamma\mu + \mu - 1 > 0$. There exists an unique equilibrium with perfect skill sorting:

---

[13]When city-specific amenities are equal to each other, $a_1 = a_2$, the indifference condition indicates that $\Psi_1(Q_1) = \Psi_2(Q_2)$, suggesting that all entrepreneur are indifferent between two cities. Essentially, two cities bring people the same real income and the same amenity. So cites are rather symmetric. This equilibrium result is conceptually unexciting. And it is not a stable equilibrium.

[14]Condition $\bar{z} > z_{min}$ stipulates that the difference in city-specific amenity, i.e. $a_1 - a_2$, should not be infinitesimally small under the perfect skill sorting equilibrium. Because in the case of $a_1 - a_2$ being infinitely close to 0, the equilibrium is essentially the symmetrical city structure result that I introduced before.

Figure 1.4: **Sorting of heterogeneous entrepreneurs between two cities.**
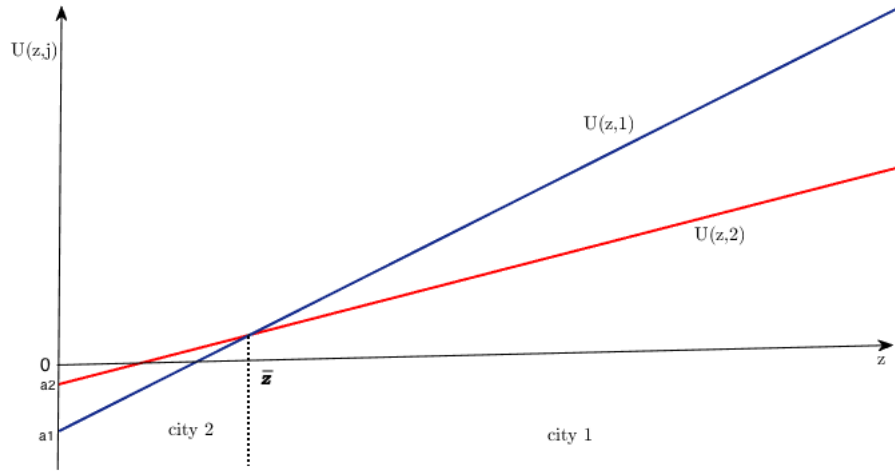
(*Note:* In this figure, the indirect utility function deriving from living in city 1 is denoted as $U(z, 1)$, and the indirect utility function deriving from living in city 2 is denoted as $U(z, 2)$. The skill sorting threshold is denoted as **z**.)

entrepreneur with initial skill higher than $\bar{z}$ choose large city, and entrepreneur with initial skill lower than $\bar{z}$ choose small city.

Figure 3.1 illustrates the skill allocation between two cities. From proposition 2, a few results immediately follow. First, the city with innovation advantage (city 1) has higher average productivity, since all the most productive entrepreneurs gather in that city. And their productivity improvement from innovation is higher due to the exogenous innovation advantage. Second, city 1 has higher nominal wage compare to city 2, because city-specific wage is positively related to city's average productivity. Third, the city with innovation advantage effectively becomes the larger (or denser) city, because workers want to live in a city with higher wage. Last, the larger city has higher housing price, because both higher wage and higher population density tend to escalate housing price. These results are summarized in the following proposition.

Proposition 3: (Equilibrium City Characteristics): Assume city 1 is endowed

with innovation advantage: $\chi_1 < \chi_2$. In equilibrium, city 1 has higher population density, $\frac{L_1}{H_1} > \frac{L_2}{H_2}$, higher average productivity, $Q_1 > Q_2$, higher wage, $w_1 > w_2$, and higher housing price, $P_{h1} > P_{h2}$.

As shown in proposition 3, this spatial sorting model can generate results that are compatible with city's cross-sectional stylized facts, such that large cities are more skilled, and have higher average wages as well as housing prices.

With perfect skill sorting, I can now write down the expressions of average city productivities:

$$
\begin{aligned}
Q_1 &= \tfrac{\alpha}{\alpha-1}\left[\tfrac{\alpha}{\alpha-1} - \left(\tfrac{\chi_1}{\gamma(\alpha-1)}\right)^{1/2}\right]\bar{z}; \\
Q_2 &= \tfrac{\alpha}{\alpha-1}\left[\tfrac{\alpha}{\alpha-1} - \left(\tfrac{\chi_2}{\gamma(\alpha-1)}\right)^{1/2}\right]z_{min}\frac{1-\left(\frac{z_{min}}{\bar{z}}\right)^{\alpha-1}}{1-\left(\frac{z_{min}}{\bar{z}}\right)^{\alpha}}.
\end{aligned}
\tag{1.4.16}
$$

From equation (4.16), we know the relative productivity is increasing in the skill sorting threshold $\bar{z}$:

$$
\frac{\partial\left(\frac{Q_1}{Q_2}(\bar{z})\right)}{\partial\bar{z}} > 0.
$$

It means that the productivity gap across cities is directly linked with skill sorting threshold [15]. If skill sorting becomes stronger over time, meaning that $\bar{z} \uparrow$ over time, the diverging trend in productivity and wage gaps across cities can be at least partially explained.

1.4.3 Top Income Inequality within Cities

So far I have calculated an entrepreneur's expected income, see equation (1.4.13). And we know that workers total income is $w_j L_j$, then top income inequality within a city can be expressed through the ratio between aggregate entrepreneurial income and labor income:

---

[15]Note, the result of $\frac{\partial\left(\frac{Q_1}{Q_2}(\bar{z})\right)}{\partial\bar{z}} > 0$ might arise due to the fat tail property of Pareto distribution. It may not stands if the productivity distribution is symmetrical.

$$entrepreneur\_share_j = \frac{\int_{i \in \Omega_j} \Pi_{ji}(z)\,dz}{w_j L_j} = \left[\frac{\gamma \alpha}{(\alpha - 1)} - 2 \left(\frac{\gamma \chi_j}{(\alpha - 1)} \right)^{1/2} \right] \frac{\bar{z}_j}{Q_j},$$

where $\bar{z}_j$ is the mean initial skill of the mass of entrepreneurs choosing city j, with $\bar{z}_j = \int_{i \in \Omega_j} z_{ji}\,di$. Taking advantage of expression (1.4.10), the entrepreneurial income share can be simplified as:

$$\Rightarrow entrepreneur\_share_j = \gamma \frac{\frac{\alpha}{(\alpha-1)} - 2\left(\frac{\chi_j}{\gamma(\alpha-1)}\right)^{1/2}}{\frac{\alpha}{(\alpha-1)} - \left(\frac{\chi_j}{\gamma(\alpha-1)}\right)^{1/2}} \qquad (1.4.17)$$

This expression suggests that a lower χ_j will raise the entrepreneurial income share, which relates the innovation activity to income distribution. Formally, this result is shown as:

Proposition 4: (Top Income Inequality within Cities): The share of entrepreneurial income within a city increases with innovation efficiency ($\frac{1}{\chi_j}$). Therefore, assumption $\chi_1 < \chi_2$ implies that big city is more unequal compared to small city.

According to this proposition, two results stand out. The first is that large city is more unequal. This result comes from the complementarity between innovation and entrepreneurial productivity. Entrepreneurs in the large city benefit more from innovation activity, because they have higher opportunities to innovate due to the city's endowed advantage. Therefore, their income share is higher even though the wage in large city is also higher.

Observing equation (1.4.17), we can see that the entrepreneurial income share in each city is completely determined by exogenous parameters: γ, α and χ_j. It suggests that top income inequality does not change over time, unless the locational innovation advantage (χ_j) changes[16].

---

[16]One way to generate increasing top income inequality is letting model adopt a different

1.5 Factors Of Increasing Skill Sorting

So far, the model has successfully delivered the right kind of sorting across cities. Now I'm going to use the comparative statics of the model to try to analyze what kind of forces might increase sorting and inequality between cities. There are two clear candidates presenting themselves in this model: TFP growth and locational fundamentals.

The aggregate TFP in this model is essentially the mean of overall entrepreneurial productivities:

$$TFP = \int_{z_{min}}^{\bar{z}} \left(\frac{\phi_2^* + \alpha - 1}{\alpha - 1} \right) z dG\left(z \right) + \int_{\bar{z}}^{\infty} \left(\frac{\phi_1^* + \alpha - 1}{\alpha - 1} \right) z dG\left(z \right) = G\left(\bar{z} \right) Q_1 + \left(1 - G\left(\bar{z} \right) \right) Q_2.$$

Since productivity follows a Pareto distribution, an increase in TFP can be proxied by an increase in the minimum productivity threshold, z_{min}, of distribution $G(\cdot)$. With perfect skill sorting, small city's average productivity Q_2 is increasing in z_{min}. But large city's average productivity Q_1 is unaffected by z_{min}, except through equilibrium sorting threshold $\bar{z}$, see equation (1.4.16). As proved in appendix A3, the equilibrium sorting threshold $\bar{z}$ increases if there is an increase in z_{min}. It implies that TFP increase will undoubtedly reinforce the skill sorting process, and increase the productivity gap across cities. This result is summarized as following:

---

innovation cost function which generates additional scale effect. For instance, if the innovation cost in city j is $\varphi\left(\phi_{ji} \mid z \right) = \frac{\chi_j w_j}{1 - \phi_{ji}} \frac{z}{Q_j}$, then the equilibrium entrepreneurial income share becomes:

$$\Rightarrow entrepreneur\_share_j = \gamma \frac{\frac{\alpha}{(\alpha-1)} - 2\left(\frac{\chi_j}{\gamma(\alpha-1)L_j} \right)^{1/2}}{\frac{\alpha}{(\alpha-1)} - \left(\frac{\chi_j}{\gamma(\alpha-1)L_j} \right)^{1/2}}.$$

The above equation clearly indicates that top income inequality is increasing in city population size L_j. Therefore, any potential forces that make the large city even large will also make it more unequal.

Figure 1.5: **Skill Sorting Becomes Stronger if TFP is Higher**
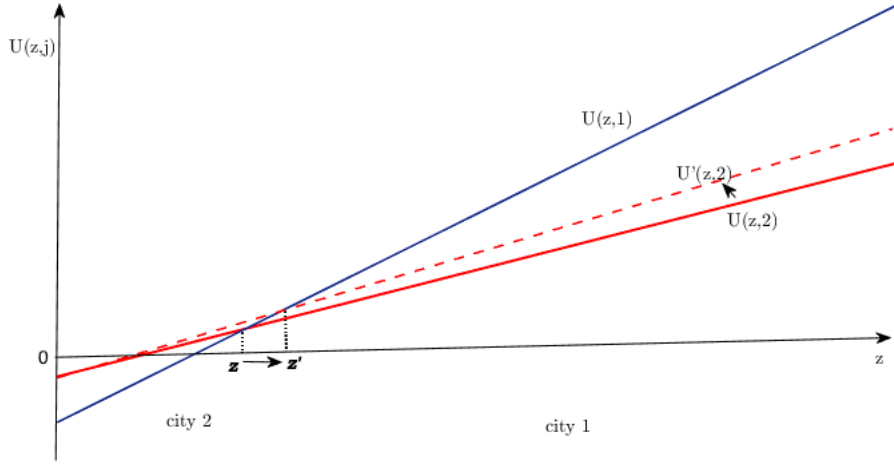
(*Note:* In this figure, the indirect utility function deriving from living in city 1 is denoted as $U(z, 1)$, and the indirect utility function deriving from living in city 2 is denoted as $U(z, 2)$. The skill sorting threshold is denoted as **z**. If $z_{min} \uparrow$, the slope of indirect utility function $U(z, 2)$ rotating to $U'(z, 2)$, which resulting threshold increase from **z** to **z'**)

Proposition 5: (Comparative Statics One): The big city disproportionately benefits from an increase in the equilibrium skill sorting threshold $\bar{z}$. Therefore, the relative productivity $\frac{Q_1}{Q_2}$, relative wage $\frac{w_1}{w_2}$, and relative population density $\frac{L_1/\bar{H}_1}{L_2/\bar{H}_2}$ between two cities are higher if TFP is higher (proxied by $z_{min} \uparrow$).

Figure 1.4.2 illustrates the effect of TFP growth on spatial skill sorting. To understand the effect of TFP growth on skill sorting, we need to examine each city closely. First of all, the rising aggregate TFP indicates that both cities are becoming more productive. Hence, both cities' average wages and housing prices will increase. However, the large city's average wage and housing price will increase more, suggesting that it will become more and more difficult for entrepreneurs being selected there. As a result, some entrepreneurs originally living in big city, who are marginally more productive than $\bar{z}$, will be better off moving to small city. Hence, the sorting threshold increases due to the endogenous responses of these marginal entrepreneurs. Meanwhile, the fat tail property of Pareto distribution

indicates that the relative productivity between the large and small cities increases whenever the sorting threshold increases. In that sense, TFP growth is equivalent to small city's technology catching up to the big city, and it strengthens the sorting process and raises the productivity gap in the end.

As for locational fundamentals, there are three factors that can affect the equilibrium sorting result: innovation advantage $(1/\chi_j)$, land supply $(1/\eta_j)$ and congestion amenity (a_j). First of all, the city endowed with innovation advantage in equilibrium becomes the large city, and it generates higher innovation opportunity, thus entrepreneurs benefit more from living in that city. Next, a city with more restricted land supply (higher η_j) has higher equilibrium housing price, which makes the less productive entrepreneurs more deterred. Third, congestion amenity (a_j) represents the downside of living in a over-crowed city, and people may not want to live in big cities if congestion amenity becomes too high to bear. The result of comparative statics *are* shown in proposition 5, and its proof can be found in appendix A3.

Proposition 6: (Comparative Statics Two): The big city disproportionately benefits from an increase in the equilibrium skill sorting threshold $\bar{z}$. Therefore, the relative productivity $\frac{Q_1}{Q_2}$, relative wage $\frac{w_1}{w_2}$, and relative population density $\frac{L_1/\bar{H}_1}{L_2/\bar{H}_2}$ between two cities are: (i) decreasing in big city's innovation advantage, $1/\chi_1$; (ii) decreasing in big city's land supply, $1/\eta_1$. In addition, both relative productivity $\frac{Q_1}{Q_2}$ and relative wage $\frac{w_1}{w_2}$ are increasing in congestion amenity, a_1, but the response of relative population density $\frac{L_1/\bar{H}_1}{L_2/\bar{H}_2}$ to an increase in a_1 is uncertain.

Figure 1.4.3 illustrates the comparative statics related to spatial skill sorting. There are two implications from the above proposition. First, if the innovation advantage in large city is declining relative to that of the small city, then skill sorting becomes stronger. This seems to be a paradoxical result, but it can be understood

Figure 1.6: **Skill Sorting Becomes Stronger if innovation Advantage declines**
or land supply more restricted or congestion amenity rises
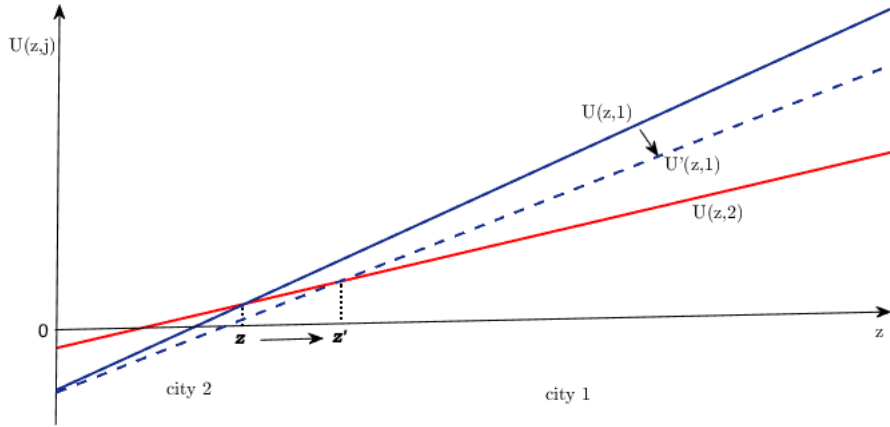
(*Note:* In this figure, the indirect utility function deriving from living in city 1 is denoted as $U(z,1)$, and the indirect utility function deriving from living in city 2 is denoted as $U(z,2)$. The skill sorting threshold is denoted as **z**. If $\frac{1}{\chi_1} \downarrow$ *or* $\frac{1}{\eta_1} \downarrow$, then the slope of indirect utility function $U(z,1)$ rotating to $U'(z,1)$, which resulting threshold increase from **z** to **z'**)

from a real world example, such as New York. One hundred years ago, the natural advantages of New York was more significant than that of Phoenix from economic point of view. Over time, with technology advancement, certain geographical and demographic advantages become less important, however, the differences in productivity and human capital composition between these two cities have increased instead of the other way around. The reason is that even though New York still has absolute advantage in innovation and production activity, but the relative advantage of New York is declining. Then the least productive entrepreneurs originally lived in New York now find themselves better off moving to cheaper cities. But, the most productive ones still stay in New York, which renders the average productivity even higher in New York. The logic behind this is: if the natural advantage is great enough, then not only the most talented, but also the "lesser" ones can make it in New York. Once again, the endogenous responses of the marginal entrepreneurs plus the fat tail property of Pareto distribution lead to a

32

larger productivity gap between New York and other smaller cities.

The second implication worth emphasizing is that spatial skill sorting becomes stronger if big city's land supply becomes more restricted. This is very intuitive, because only the richest and the most talented people can afford high housing price. Again, we can understand it from a real world example. As the innovation cluster center in twenty-first century, Silicon valley is currently among the most expensive city in United States. However, the city's high housing prices reflect more than just good weather and high incomes. The city has rather severe restrictions on home construction. Between 2001 and 2008, despite the booming demand, the area's stock of single-family homes increased by less than 5%, which was less than 1/3 of the U.S. average building rate over that period. According to Glaeser (2011), Silicon valley's housing price would be 40% lower if there is no restriction on house/land supply.

1.6 Localized Knowledge Spillover

So far, the baseline model assumes that one city has exogenous innovation advantage, which attracts the most talented entrepreneurs living there. And these entrepreneurs in turn push up this city's wage and housing price, making it too expensive for the relatively less talented entrepreneurs. As introduced in the abstract, one of the main contributions of this paper is to build a model that allowing for endogenous innovation advantage. This section is to formalize that extension through the effect of localized knowledge spillover. According to Lucas and Moll (2014), entrepreneurs can interact with and learn from other entrepreneurs in the same place. On top of their idea, I add the spatial aspect which allowing geography to play a role in the process of idea exchange or innovation. Essentially, innovation is more likely to happen if entrepreneurs live in the same city with other top talented entrepreneurs, since face-to-face meeting is the more effective

form of human interaction. This extension has the following three merits: (i) it shows that big cities can be endogenously more innovation and more productive without assuming any exogenous fundamental differences across different cities; (ii) it also allows me to formally analyze how information technology affects the skill compositions at different cities; (iii) its modeling approach is very similar to that of the baseline model, hence it preserves the results of the comparative statics from the baseline model.

To present the idea at its simplest, the innovation technology is similar to that of the baseline model, except now innovation depends on an endogenous variable: local learning opportunities (β_j). The innovation process is as following. An entrepreneur i in city j can buy a chance of innovation with probability ϕ_{ji}, however only with chance β_j, innovation in city j can actually occur. With probability $(1 - \phi_{ji}\beta_j)$ his productivity remains unchanged, and the ones obtaining a chance to innovate draws a new skill z^+ from distribution $G(\cdot)$. The expected productivity level conditional on innovation is the same as baseline model:

$$E\left(z^+ \mid z, \ innovation\right) = \int_z^\infty x dG(x) = \frac{\alpha z}{\alpha - 1};$$

the added "plus" superscript refers to the productivity after the innovation decision. The expected productivity for innovation probability ϕ_{ji} is:

$$\tilde{q_{ji}}\left(z\right) = \underbrace{\phi_{ji}\beta_j\frac{\alpha z}{\alpha - 1}}_{expected \ productivity \ if \ innovate} + \underbrace{\left(1 - \phi_{ji}\beta_j\right) z}_{innovation \ not \ occur} \qquad (1.6.1)$$

$$\implies \tilde{q_{ji}}\left(z\right) = \left(\tfrac{\phi_{ji}\beta_j + \alpha - 1}{\alpha - 1}\right) z.$$

The variable β_j stands for local chance of meeting people with higher productivity. Formally, local learning environment is characterized by a function $\beta_j = \beta(\bar{z}_j)$, and $\beta(\bar{z}_j) = \{\bar{z}_j \in (0, \infty) : \ \beta(\bar{z}_j) \in (0, 1)\}$, where $\bar{z}_j = \int_{i \in \Omega_j} z_{ji} di$ representing the average initial productivity of entrepreneurs living in city j. To introduce the

learning technology in detail, the following assumption is necessary:

Assumption: $\beta(\cdot)$ is continuous, concave, and increasing in the average initial productivity of entrepreneurs, $\bar{z}_j$, who endogenously choose city j. Meanwhile it has following properties: $\beta(0) = 0$ and $\beta(\infty) = 1$.

The above assumption indicates that as a city becomes more skilled, the chance of its residents improving productivity is increasing as well. In addition, $\beta(\bar{z}_j)$ increases faster as $\bar{z}_j$ is smaller, eventually it slowly approaches 1 when $\bar{z}_j$ approaches infinity. There are many functions satisfy the above assumption, for analysis purpose, I choose the following form:

$$\beta(\bar{z}_j) = \frac{\bar{z}_j}{\bar{z}_j + c}, \quad c > 0; \tag{1.6.2}$$

The production process in this section is the same as the baseline model, indicating that all the results stated in proposition 1 in baseline model are also valid here. Therefore, the expected profit from selling variety remains the same: $\tilde{E}\left(\pi_{ji} \mid z\right) = \frac{\gamma w_j L_j}{Q_j}\tilde{q}_{ji}(z)$. However, the baseline model assumes different innovation costs at different locations. The extended model is designed to eliminate that exogeneity, since it might be mixed with the endogenous effect of localized idea exchange. Therefore, I assume there is no fundamental differences in city's innovation costs:

$$\tilde{\varphi}\left(\phi_{ji} \mid z\right) = \frac{\chi z}{1 - \phi_{ji}}\frac{w_j L_j}{Q_j}, \quad \chi > 0; \tag{1.6.3}$$

The process of solving for the optimal innovation opportunity is similar to the baseline model. Its expression is:

$$\tilde{\phi}_{ji} = 1 - \left(\frac{\chi(\alpha - 1)}{\gamma \beta_j}\right)^{1/2} = \tilde{\phi}_j \tag{1.6.4}$$

Clearly, if a city provides a better chance to meet more productive people,

namely a higher β_j, then the equilibrium innovation opportunity $\tilde{\phi}_j$ will be higher, which leads to higher return of innovation investment. Substituting $\tilde{\phi}_{ji}$ in (1.5.1) with above equation, we can get city's average productivity:

$$Q_j = \int_{\Omega_j} \left(\frac{\tilde{\phi}_j \beta(\bar{z}_j) + \alpha - 1}{\alpha - 1} \right) z_{ji} di = \underbrace{\left[\frac{\alpha - 1 + \beta(\bar{z}_j)}{\alpha - 1} - \left(\frac{\chi \beta(\bar{z}_j)}{\gamma (\alpha - 1)} \right)^{1/2} \right]}_{Scale\ of\ improvement\ due\ to\ innovation} \bar{z}_j.$$

(1.6.5)

The average productivity in a city Q_j depends on two factors: the average productivity of entrepreneurs living in that city, $\bar{z}_j$, and the scale of productivity improvement from innovation. Furthermore, the scale of improvement is complement to the level of $\bar{z}_j$, indicating that there exists a complementarity between innovation and entrepreneurial productivity. It is precisely this complementarity that generates the result of spatial skill sorting.

In this section, I will simplify the process of skill sorting due to the similarity to baseline model. Therefore, I now jump ahead and introduce the expression of entrepreneur's indirect utility function for living in city j:

$$\tilde{U}_i^e(j, z) = \frac{\tilde{\Gamma}_j}{P_{hj}^{1-\mu}} z - a_j = \tilde{\Psi}_j (Q_j) z - a_j.$$

(1.6.6)

where $\tilde{\Psi}_j (Q_j) = \frac{\gamma}{\eta_j} \left(\frac{\theta}{u^* + a_j} \right)^{\frac{\mu}{1-\mu}} \left[\frac{\alpha - 1 + \beta(\bar{z}_j)}{(\alpha - 1)} - 2 \left(\frac{\chi \beta(\bar{z}_j)}{\gamma(\alpha - 1)} \right)^{\frac{1}{2}} \right] Q_j^{\frac{\gamma \mu + \mu - 1}{1-\mu}}$ represents the real benefit of living in city j, and it is the slope of indirect utility function. Clearly, as long as condition $\frac{\gamma \mu + \mu - 1}{1-\mu} \geq 0$ holds, there is $\frac{\partial \tilde{\Psi}_j(Q_j)}{\partial Q_j} > 0$. This condition means that the benefit of living in an increasingly more productive city outweighs its increasing housing cost, suggesting that more productive entrepreneurs benefit more from big cities.

With indirect utility function settled, entrepreneurs choose the optimal location by comparing their utilities of living at different cities. Once again, the linear

36

relationship between indirect utility function $\tilde{U}_i^e(j, z)$ and initial skill z indicates perfect skill sorting. The equilibrium analysis is more complicated compare to the baseline model, because there exists the possibility of multiple equilibria due to the fact that entrepreneur's location choice now depending on other's choices as well. The first possible equilibrium is that all entrepreneurs with productivity higher than certain threshold $\bar{z}^*$ choosing city 1, and the second possible equilibrium is that all entrepreneurs with productivity higher than certain threshold $\bar{z}^*$ choosing city 2. Since city is generic and has no fundamental differences except for city-specific amenity, a_j, which representing the degree of congestion or crowdedness. I can assign one of the city to be the more crowded city, then that city in equilibrium will become the denser city. To be consistent with baseline model's notation, I let the first city be the more crowded one, such that $a_1 > a_2$. Based on this assumption, there exists an unique skill sorting threshold $\bar{z}^*$, such that entrepreneurs with initial skill higher than $\bar{z}^*$ choose city 1. The threshold $\bar{z}^*$ is:

$$\bar{z}^* = \frac{a_1 - a_2}{\tilde{\Psi}_1\left(Q_1\right) - \tilde{\Psi}_2\left(Q_2\right)}. \tag{1.6.7}$$

If this threshold is larger than the minimum level z_{min} of the initial productivity distribution $G(\cdot)$, $\bar{z}^* > z_{min}$, the spatial equilibrium characterized by perfect skill sorting is as stated in proposition 7:

Proposition 7: (Localized Knowledge Spillover and Skill Sorting) Assume $a_1 > a_2$ and $\gamma\mu + \mu - 1 > 0$. With localized knowledge spillover, there exists an unique equilibrium with perfect skill sorting: entrepreneur with initial skill higher than $\bar{z}^*$ choose large city, and entrepreneur with initial skill lower than $\bar{z}^*$ choose small city.

37

1.6.1 The Effects of Information Technology

So far, localized knowledge spillover alone still can generate the result that agents with different skills are sorted into cities with different productivities. And the reason for this result is that more talented entrepreneurs benefit more from big cities due to the learning technology. The next step is to analyze how changes in information technology affect the pattern of skill sorting. To implement this analysis, I let $c \downarrow$ representing the improvement in information or communication technology, because equation (1.6.2) indicates that $\frac{\partial \beta(\bar{z}_j)}{\partial c} < 0$. Intuitively, modern technology improves people's ability to interact or communicate with each other. It means that people's chance to exchange ideas becomes stronger as c decreases. This functional form is mathematically easy, but it is able to incorporate the effects of information technology on innovation. This line of study is particularly interesting, because the way people interact with each other have changed drastically in modern society. For example, emails and video chats have infiltrated almost every aspect of day-to-day life and business operations as well. I want to explore how these improvements in information technology and telecommunication change the value of cities, and how they affect the productivity gap between cities. The key analysis is to know the effect on relative productivity, which is expressed as following:

$$\frac{Q_1}{Q_2} = \underbrace{\frac{\frac{\alpha-1+\beta(\bar{z}_1|c)}{(\alpha-1)} - \left(\frac{\chi\beta(\bar{z}_1|c)}{\gamma(\alpha-1)}\right)^{\frac{1}{2}}}{\frac{\alpha-1+\beta(\bar{z}_2|c)}{(\alpha-1)} - \left(\frac{\chi\beta(\bar{z}_2|c)}{\gamma(\alpha-1)}\right)^{\frac{1}{2}}}}_{Direct\ effect} \cdot \underbrace{\frac{\frac{\bar{z}^*}{z_{min}} - \left(\frac{z_{min}}{\bar{z}}\right)^{\alpha-1}}{1 - \left(\frac{z_{min}}{\bar{z}}\right)^{\alpha-1}}}_{Sorting\ effect} \tag{1.6.8}$$

Information technology improvements generate two different effects on relative productivity: the "direct effect" represents the relative improvement of learning environment in different cities, and the "sorting effect" represents the relative

change of skill composition due to spatial sorting. As discussed above, the advancement in information technology ($c \downarrow$) improves city's learning environment, meaning $\frac{\partial \beta_j}{\partial c} < 0$. On the other hand, depending on the skill composition of each city's, $\bar{z}_j$, a decline in c will increase $\beta(\bar{z}_j)$ to different degrees, and this differential effect is determined by the sign of $\frac{\partial^2 \beta(\bar{z}_j)}{\partial Q_j \partial c}$. Depending on whether information technology is complement or substitute with city, I summarize the effects in two cases. According to Gaspar and Glaeser (1996), when telecommunications technology improves, there are two opposing effects on cities and face-to-face interactions: some relationships that used to be face-to-face will be done electronically (an intuitive substitution effect), and some individuals will choose to make more contacts, many of which result in face-to-face interactions. In the case of them being complement, Silicon Valley is a good example. People in Silicon Valley usually only require two things to work, phone and computer, so they can easily connect electronically. They could have worked from anywhere, yet they choose to live in the most expensive city in U.S.. It implies that human interactions are essential, and cannot be replaced by telecommunications.

(i) The condition for information technology and city being complement is: $0 < Q_2 < Q_1 < c$[17]. Whenever information technology and city are complement to each other, information technology advancement will improve big city's learning opportunity to a greater degree. In this case, whenever there is a drop in c, the increase in the real benefit of living in big city will be greater than the increase in small city, such that $\triangle \tilde{\Psi}_1(Q_1) > \triangle \tilde{\Psi}_2(Q_2)$. According to equation (1.5.8), the equilibrium sorting threshold $\bar{z}^*$ will decrease in this scenario. Intuitively, if there is a greater improvement in big city's learning opportunity as a result of information technology advancement, then the less productive entrepreneurs will

---

[17]The technical condition for information technology and city being complement is as following: whenever $0 < Q_2 < Q_1 < c$, then $\frac{\partial^2 \beta(\bar{z}_j)}{\partial Q_j \partial c} < 0$, which implies that a decline in c will lead to a greater increase in β_1 than β_2, such that $\triangle \beta(\bar{z}_1 \mid \nabla c) > \triangle \beta(\bar{z}_2 \mid \nabla c)$.

be tempted to move into big city for its greater benefits despite the high housing price, hence $\bar{z}^* \downarrow$. These two different forces will lead to two competing effects on city's relative productivity:

$$\frac{Q_1}{Q_2} = \underbrace{\frac{\frac{\alpha-1+\beta(\bar{z}_1|c)}{(\alpha-1)} - \left(\frac{\chi\beta(\bar{z}_1|c)}{\gamma(\alpha-1)}\right)^{\frac{1}{2}}}{\frac{\alpha-1+\beta(\bar{z}_2|c)}{(\alpha-1)} - \left(\frac{\chi\beta(\bar{z}_2|c)}{\gamma(\alpha-1)}\right)^{\frac{1}{2}}}}_{Direct\ effect} \uparrow \cdot \underbrace{\frac{\frac{\bar{z}^*}{z_{min}} - \left(\frac{z_{min}}{\bar{z}^*}\right)^{\alpha-1}}{1 - \left(\frac{z_{min}}{\bar{z}^*}\right)^{\alpha-1}}}_{Sorting\ effect} \downarrow;$$

The increasing effect of ∇c on the "direct effect" captures the relatively greater change in big city's learning environment; the decreasing effect of $\bar{z}^* \downarrow$ on the "sorting effect" captures the negative effect on skill sorting. If the negative sorting effect in part B is the dominant force, then the equilibrium relative productivity decreases as information technology improves. On the other hand, if the positive effect in part A is the dominant force, then the equilibrium relative productivity increases as information technology improves.

(ii) The condition for information technology and city being substitute is: $0 < Q_2 < c < Q_1$ or $0 < c < Q_2 < Q_1$[18]. Whenever information technology and city are substitute to each other, information technology advancement will improve small city's learning opportunity to a greater degree. In this case, the increase in the benefit of living in a big city will be smaller than the increase in small city, then the value of small city inherently rises. Therefore, small cities become more attractive since the barrier of knowledge spread now declines. According to equation (1.5.8), the equilibrium sorting threshold $\bar{z}^*$ will increase accordingly. Intuitively, if the improvement of learning environment in big city is smaller compare to small city, then the marginally less productive entrepreneurs, who originally live in big city, will move to small city instead, hence $\bar{z}^* \uparrow$. As a consequences of these two different forces, there will be two competing effects on city's relative productivity:

---

[18]The technical condition for information technology and city being substitute is as following: whenever $0 < Q_2 < c < Q_1$ or $0 < c < Q_2 < Q_1$, then $\frac{\partial^2 \beta(\bar{z}_j)}{\partial Q_j \partial c} > 0$, which implies that a decline in c will lead to a greater increase in β_2, such that $\triangle\beta(\bar{z}_1 \mid \nabla c) < \triangle\beta(\bar{z}_2 \mid \nabla c)$.

$$\frac{Q_1}{Q_2} = \underbrace{\frac{\frac{\alpha-1+\beta(\bar{z}_1|c)}{(\alpha-1)} - \left(\frac{\chi\beta(\bar{z}_1|c)}{\gamma(\alpha-1)}\right)^{\frac{1}{2}}}{\frac{\alpha-1+\beta(\bar{z}_2|c)}{(\alpha-1)} - \left(\frac{\chi\beta(\bar{z}_2|c)}{\gamma(\alpha-1)}\right)^{\frac{1}{2}}}}_{part\ A} \downarrow \cdot \underbrace{\frac{\frac{\bar{z}^*}{z_{min}} - \left(\frac{z_{min}}{\bar{z}^*}\right)^{\alpha-1}}{1 - \left(\frac{z_{min}}{\bar{z}^*}\right)^{\alpha-1}}}_{part\ B} \uparrow;$$

The decreasing effect of ∇c on the "direct effect" captures the relatively smaller change in big city's learning environment; the increasing effect of $\bar{z}^* \downarrow$ on the "sorting effect" captures the positive effect on the relative skill composition at two cities. Like the first case, there is no way to determine which part is the dominant effect.

Now it is time to summarize the above effects. When telecommunication and city are complementary, the benefits of big cities increased more than that of small cities, because information technology has made it more and more valuable for people to stay close to the top talents, thus big cities will attract more and more relatively unskilled people. The opposite tradeoff occurs when telecommunication and city are substitute. In that case, the benefits of small cities increased relatively more, because information technology has rendered it easier for people to learn from top talents over long distance, thus small cities will attract more and more relatively unskilled people.

Essentially, the effect of localized knowledge exchange provides another perspective as to why large cities have better environment for idea exchange, without assuming exogenous differences on locational fundamentals. Therefore, when innovation advantage in large city is shut down, meaning that $\chi_1 = \chi_2$, localized knowledge spillover alone still can produce perfect skill sorting result. The reasoning behind this is that large cities are not only the places with dense population, they are also the places where the best and brightest minds live. Knowledge spread is simply faster or more effective when people live close to the ones that with the best ideas. Just as Marshall (1890) described how in dense concentrations "the mysteries of the trade become no mystery but are, as it were, in the air." To simply

put, hanging around successful people will improve the chance of people becoming successful themselves. Everyone would want to live in the city with best learning opportunities. But the most talented entrepreneurs are those most able to take advantage of these opportunities and so most willing to pay for them.

1.7 Model Extension: Trade

Economic activity depends crucially on the transportation of goods and people across space. So far the model focuses on the interaction of cities through the mobility of people. This section explores how the input-output linkages affect spatial concentration through including bilateral trade into the baseline model. Furthermore, trade and transportation cost have declined greatly over time. It would be interesting to analyze how the decrease in trade cost shapes economic activity across space.

Bilateral trade occurs both at the differentiated intermediate goods and final consumption goods level. Trade cost is assumed to be the typical iceberg form, meaning $\tau_{nj} \geq 1$ units of must be shipped from city n in order for one unit to arrive in region j. It is necessary to assume asymmetrical trade cost for the purpose of forming asymmetric cities. The asymmetric trade cost might arise from a number of considerations, such as land gradient and trade volume (per-unit iceberg trade cost is lower when trade volume is larger). In addition, big cities, due to scale economy, might provide better or less-costly services related to trade such as insurance, which then leads to lower ice-berg cost. There are other plausible channels that capable of generating trade asymmetry. However, it is theoretically inconsistent to analyze such channels under the framework of this paper. For instance, Waugh (2010) believed that the asymmetric trade volume is highly correlated with income level, rather than geographic distance. And it is well known that the consumption baskets of high and low-income consumers look

very different (e.g., Deaton and Muellbauer, 1980). Fajgelbaum and Khandelwal (2016) suggested that the asymmetry in trade hinges on the fact that poor consumers spend relatively more on tradable sectors, while high-income individuals consume relatively more services, which are the least traded sector. Therefore, trade balance condition will hold because there exists a lower rate of substitution between imports and domestic goods for the relatively poor. As we can see, the above methods for modeling asymmetric trade system focus on demand side, and it requires two sectors: tradable and non-tradable. But it is conceptually difficult to adopt such extra structure under this framework. Whereas asymmetric trade cost assumption is a relatively easy approach in this paper.

The production process is the same as the baseline model. In each city there is a final good producer that supplies the final output good competitively. And final good serves as numeraire in both cities. Entrepreneurs produce differentiated intermediate goods and have claims on the net profit. The only difference between baseline model and the extended model is that the production of final good in a city requires intermediate goods input from both cities instead of just local varieties. The final good output in city j is defined as following:

$$Y_j = \left(\sum_{n \in \{1,2\}} \int_{i \in \Omega} y_{nji}^\theta di \right)^{1/\theta}, \qquad \theta \in (0,1) \tag{1.7.1}$$

where y_{nji} is the amount of intermediate good i (i indexing for entrepreneur) used for final good production in city j shipped from city n. Profit maximization implies that intermediate good price is a constant markup over the marginal cost of supplying a variety,

$$p_{nji} = \frac{1}{\theta} \frac{\tau_{nj} w_n}{q_{nji}^\gamma}. \tag{1.7.2}$$

Due to the love of variety, all entrepreneurs sell their differentiated intermediate

goods to both cities. The final output good is numeraire in each city, then the demand of intermediate good i from city n to city j is

$$y_{nji} = \left(\frac{\tau_{nj}w_n}{\theta q_{nji}^\gamma}\right)^{\frac{1}{\theta-1}} Y_j.$$

(1.7.3)

Factor market clearing implies that local aggregate output Y_j is the sum of imports from all locations. Replacing (1.5.3) into final good production function (1.5.1), the factor market clearance condition is:

$$\left[\left(\frac{\theta}{w_j}\right)^{\frac{1}{\gamma}} Q_j Y_j + \left(\frac{\theta}{\tau_{nj}w_n}\right)^{\frac{1}{\gamma}} Q_n Y_j\right] = Y_j.$$

(1.7.4)

With intermediate goods price function (1.6.2) and demand function (1.6.3), the value of location$j's$ imports from location n can be expressed as:

$$X(n,j) = \int p_{nji}y_{nji}di = \left(\frac{\theta}{\tau_{nj}w_n}\right)^{\frac{1}{\gamma}} Q_n Y_j.$$

(1.7.5)

Where Q_j still represents the average productivity of entrepreneurs living in city j, and its definition is $Q_j = \int_{\Omega_j} \left(\frac{\phi_{ji}+\alpha-1}{\alpha-1}\right) z_{ji}di$. Trade volume expressed in (1.6.5) states that the volume of city j's imports from city n depends on a bilateral trade friction τ_{nj}, an origin-specific cost (inverse) term $\left(\frac{\theta}{w_n}\right)^{\frac{1}{\gamma}} Q_n$, and target market's size Y_j. Goods market clearing and balanced trade imply that for city $n, j \in \{1, 2\}$, $n \neq j$:

$$X(n,j) = X(j,n).$$

Replace $X(n,j)$ with (1.6.5) into the trade balance condition, we can get the following result:

$$\left(\frac{\theta}{\tau_{12}w_1}\right)^{\frac{1}{\gamma}} Q_1 Y_2 = \left(\frac{\theta}{\tau_{21}w_2}\right)^{\frac{1}{\gamma}} Q_2 Y_1.$$

(1.7.6)

The spatial equilibrium is the same as baseline mode, with an additional condition for balanced trade. Next I show how the process of innovation and spatial sorting changes.

1.7.1 Innovation and Skill Sorting

Entrepreneurs face the same choices in the extended trade model. First they decide where to stay. Then they decide the optimal investment on innovation based on initial skill as well as the location choice. After which, production takes place. There is no decision needed on whether or not to export, since each and every one of the entrepreneurs will export. Using equations (1.6.2) and (1.6.3), the aggregate profit for entrepreneur i located at city j before innovation cost is

$$E\left(\pi_{ji}^{X} \mid z\right) = (1-\theta)\left(\frac{\theta}{w_j}\right)^{1/\gamma}\left[Y_j + \left(\frac{1}{\tau_{jn}}\right)^{1/\gamma} Y_n\right]\left(\frac{\phi_{ji} + \alpha - 1}{\alpha - 1}\right) z. \qquad (1.7.7)$$

The expected profit in (1.6.7) indicates that if the cost of shipping goods from two cities are different, or in this case large city is cost efficient $\tau_{12} < \tau_{21}$, then choosing which city to stay has direct effect on the profit from trade. Using the trade balance condition (1.6.6) and factor market clearance condition (1.6.4), the expected profit function can be simplified as

$$E\left(\pi_{ji}^{X} \mid z\right) = (1-\theta)\frac{Y_j}{Q_j}\left(\frac{\phi_{ji} + \alpha - 1}{\alpha - 1}\right) z. \qquad (1.7.8)$$

Comparing the expected profit function (1.4.5) from the baseline model with expression (1.6.8), we will notice that they are of the exact same form. This observation indicates the equilibrium innovation opportunity for each entrepreneur under the two models are exactly the same. Peculiar as this result may seem, it makes sense mathematically. Because both the expected profit function and the

innovation cost function are linear to city's economy size Y_j and entrepreneur's expected productivity q_{ji}. This linearity relationship along the trade balance condition indicate the equilibrium innovation opportunity ϕ_{ji}^X is

$$\phi_{ji}^X = 1 - \left(\frac{\chi_j(\alpha-1)}{\gamma}\right)^{1/2} = \phi_j^X \qquad (1.7.9)$$

With the expression of expected profit in (1.6.8) and the optimal innovation opportunity in (6.9), the expected net profit is

$$\Pi_{ji}^X(z) = \Gamma_j^X z. \qquad (1.7.10)$$

where $\Gamma_j^X = \left[\frac{\gamma\alpha}{(\alpha-1)} - 2\left(\frac{\gamma\chi_j}{(\alpha-1)}\right)^{1/2}\right]\frac{w_j L_j}{Q_j}$ summarizes entrepreneur's return of productivity for living in city j, which includes the gain from both innovation and trade. With expected net profit , the skill sorting process is determined by the following utility maximization problem:

$$\max_{j\in\{1,2\}} U(z,j) = \Psi_j^X z - a_j. \qquad (1.7.11)$$

where $\Psi_j^X = \left(\frac{1}{\eta_j}\right)^{1-\mu}\left[\frac{\gamma\alpha}{(\alpha-1)} - 2\left(\frac{\gamma\chi_j}{(\alpha-1)}\right)^{1/2}\right]\frac{(w_j L_j)^\mu}{Q_j}$ represents the real benefit of entrepreneur living in city j. The process of skill sorting among entrepreneurs is the same before, which means there exists an unique equilibrium skill threshold $\bar{z}^X$, such that entrepreneurs with initial skills higher than $\bar{z}^X$ choosing to live in the big city, and entrepreneurs with initial skills lower than $\bar{z}^X$ choosing to live in small city. The threshold $\bar{z}^X$ is defined by the following equation:

$$\bar{z}^X = \frac{a_1 - a_2}{\Psi_1^X - \Psi_2^X}. \qquad (1.7.12)$$

As proved in appendix A5, when trade cost is asymmetrical, with the large city facing lower trade cost, skill sorting will become stronger as trade costs decline.

For example, if both city's ice-berg trade costs decline by the same percentage, say 50%, then the sorting threshold increases accordingly. Proposition 8 captures this result.

Proposition 8: (Trade Cost Decline and Skill Sorting) With asymmetrical trade costs, $\tau_{21} > \tau_{12}$, there exists an unique equilibrium with perfect skill sorting: entrepreneur with initial skill higher than $\bar{z}^X$ choose large city, and entrepreneur with initial skill lower than $\bar{z}^*$ choose small city. In addition, the relative productivity $\frac{Q_1}{Q_2}$ and relative wage $\frac{w_1}{w_2}$ across cities become higher as trade costs decline by the same proportion.

In a economy with asymmetric trade costs, entrepreneurs in large city benefit more from a trade cost decline, because they end up spending less on shipping goods to the other city, see equation (1.6.2). In equilibrium, the benefits of lower trade cost in large city leads to stronger competition among entrepreneurs. As a result, the relative wage between large and small city $\frac{w_1}{w_2}$ increases as trade costs decrease, see detailed proof in appendix A5. In equilibrium, spatial skill sorting becomes stronger in response to trade cost decline, and the productivity as well as wage gaps across cities become wider.

1.8 Conclusions

This paper builds a model to integrate innovation, skill sorting and agglomeration in a multi-city framework. In the model, innovation and knowledge spillover are the main forces of agglomeration. The fundamental difference between cities is that some city has better environment in stimulating innovation, which is reflected as differences in learning opportunities. Everyone would like to be where learning opportunities are the greatest, so that they can improve productivity. But only the most talented entrepreneurs rise above the competition, because

they are the ones who can afford the higher wages and higher housing prices. In large cities, entrepreneurs invest more resources to reap the benefits of innovation, hence innovation activities heavily concentrate in these cities. Meanwhile, the complementarity between skill sorting and innovation causes the large cities having higher productivity comparing to small cities. The productivity premium in large city then leads to higher wages. In equilibrium, the higher nominal wages compensate the higher living costs in big cities.

This paper focus on the interaction between innovation concentration, skill concentration, and their effects on spacial inequality. Cities play an active role in this framework. A few city-specific factors serve as channels for spatial skill sorting, such as innovation advantage, housing supply, and congestion amenity. Through comparative static analysis, we can get some insights on what are the driving forces behind the increased skill sorting over the past three decades. In addition, the paper provides a theory on how TFP growth and trade cost decline affect the distribution of human capital across cities. Over all, this framework prove to be tractable enough to explore what is causing the great divergence among American cities.

Furthermore, the extended model in section 1.3 discussed the effects of localized knowledge spillover. Spatial knowledge spillover has long been perceived as one of the most important causes for urban agglomeration. Many theories have proved that sharing knowledge or skill through social interaction can generate significant spillover effect, which is considered as the key to economic growth. The emphasis of this paper is to provide a simple theory of how localized knowledge spillover affects the geographical divergence across cities. In larger cities, the environment of idea exchange of knowledge spread is better, because they are the places where the most talented individuals live. And every one wants to hang out with the smart or successful people, since it gives them the best chances to improve productivities

and enhance innovations. Thus the most skilled and most expensive cities attract more talented individuals, and become more skill-intensive. This will ultimately lead the American cities on a diverging trend. And the fact has proved this point. Cities like New York, San Francisco and Seattle are not bogged down by high living costs, they are in fact on the rise compared to the rest of cities.

Chapter 2

# Household Savings Behavior And Coresidence In Urban China

2.1 Introduction

Since the implementation of housing privatization reform in 1998, there has been significant housing price appreciation in China. The average house value-to-income ratio in mega-cities such as Beijing and Shanghai has been over 10 since 2010, and it is is about 6 or higher in second and third tier cities. As banks in China impose 30% minimum down payment, the house price-to-income ratio indicates it would take about two years of household income to make a down payment. Suppose household took a mortgage loan for the other 70% of the house price, with 6% mortgage interest rate and 20 years maximum maturity, it would take 40% of annual household income to service the mortgage loan. Therefore, it is too heavy a burden for the average household to buy a house. The natural question to ask is how are Chinese people affording the outrageous housing price? This paper examines the relationship between the rising house prices and household saving rate, and how households solve the housing affordability issues through intergenerational transfer and coresidence.

The Chinese household savings rate has risen from 18% in 1998 to 32% in 2012. This persistent increase in savings rate is extraordinary even among the fast-growing economies. Aside from the rising household savings rate across all board, the savings rate of the young households is particularly high, even higher than

the middle-aged, see figure 2.1. This age-savings profile is difficult to understand within a standard life cycle model. In a conventional life cycle model, when income growth trend is high, the young should save less than the middle-aged in anticipation of future income prospects. One way to reconcile this inconsistency is incorporating housing choices into the life cycle model. There are two living arrangements most common in urban China, coresidence and buying a home, both having implications on age-savings pattern[1]. When the credit market is underdeveloped, with the strong motivation for buying homes, young adults are forced to reduce consumption in facilitating savings during the early periods. On the other hand, shared housing between adult children and parents implies the aggregation of savings within households, hence household savings will be affected by coresidence choice.

Figure 2.1: **Life-cycle profile on savings rate in 2012**
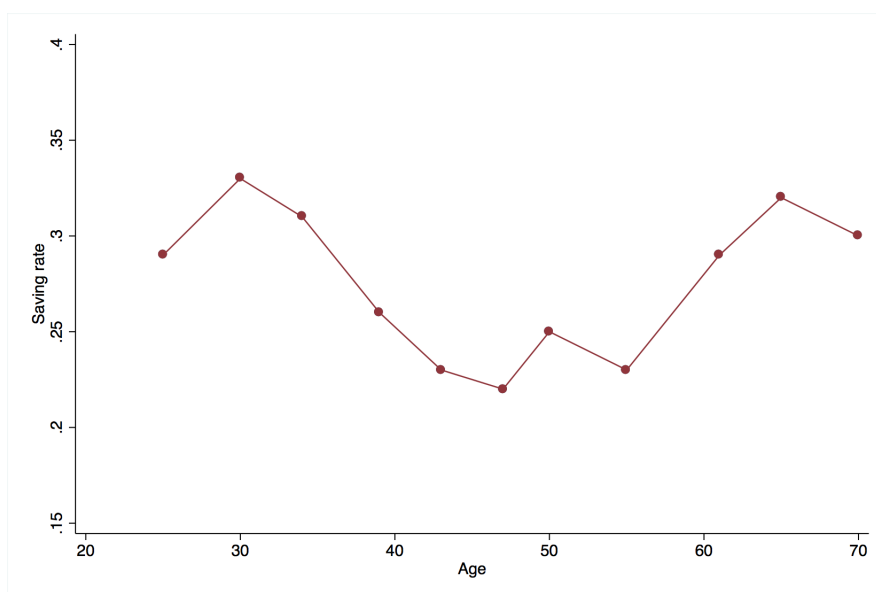
Note: the age-savings profile of 2012 is calculated from China Household Finance Survey (CHFS).

---

[1] Due to the nascent nature of housing rental market, only 15% households rent, among which, two thirds are public rental housing. Hence renting is not considered in this paper.

In a fast growing economy, intergenerational coresidence would be expected to decline with socio-economic development, because extended family ties will be weakened by the increasing labor opportunities and growing income prospects. But the evidence of intergenerational coresidence from China Household Finance Survey in 2012 showed otherwise. Coresidence in urban China is quite common, in fact, about 45% young adults aged 35 still live with at least one of their parents, see figure 2.2. The coresidence pattern in China reflects that the younger generation is very much dependent on the housing support of the older generation. Due to the housing privatization reform in the 1990s, 90% of urban households bought their own houses at cheap prices, thus housing is not an issue for the older generation. The younger generation, however, has to face severe financial burden caused by the high housing prices. Therefore, coresidence is an important mechanism for young adults to lower consumption and generate higher savings. In addition, the age-savings profiles can be separately examined based on different household structures: coresidence and non-coresidence.

Aside from coresidence, intergenerational transfer is another important mechanism for young adults to afford expensive housing. According to the China Health and Retirement Longitudinal Study (CHARLS) survey, 13% parents buy a house for their children as wedding gifts. And more than 70% of parents give money transfer to their children during the wedding event, where the average value is about 3 times of the annual household income. To realize this level of intergenerational transfer, parents need to have substantial savings. It suggests that the financial burden of housing purchase not only affects young household's saving and living situation, it also bears significance on older household's saving decisions.

To fix ideas, I built a life-cycle dynastic model with endogenous housing decision and intergenerational transfer. The model incorporates two generations: old

Figure 2.2: **Comparison of life-cycle coresidence patterns in urban China and U.S.**
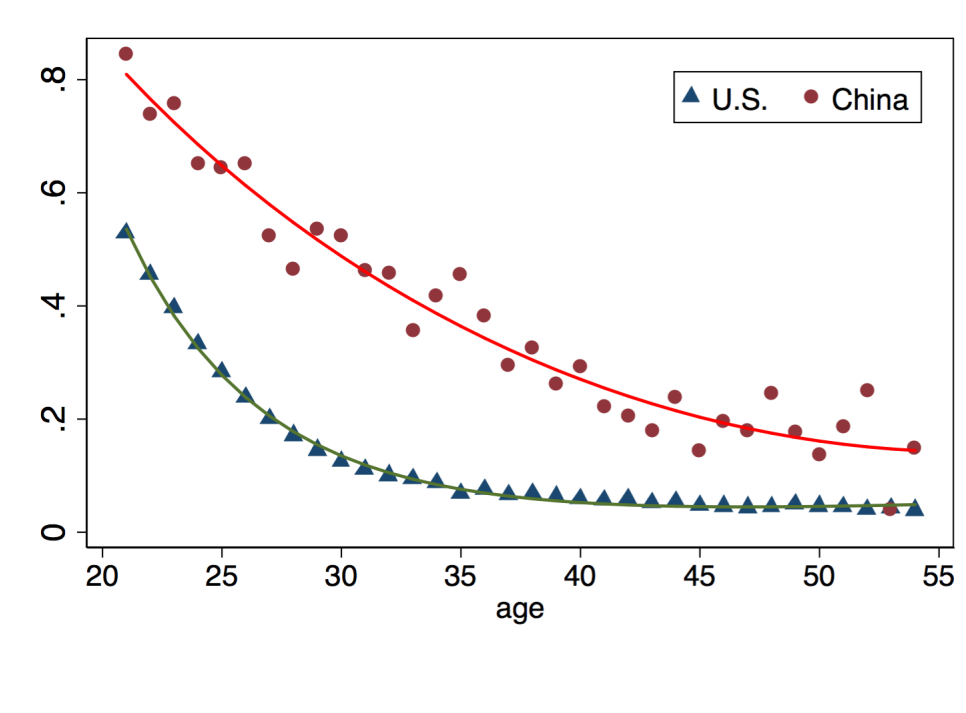
Note: people who are at school are not included when calculating the age-coresidence pattern.

parents who live from 45 to 80 years old, and adult children from 20 to 80. In line with the extremely high homeownership caused by housing privatization reform in the 1990s, the older generation are home owners to begin with, and adult children make the optimal living choice at each period: owning a house or living with parents. Coresidence promises adult children free housing but bears disutility; once adult children buy their own houses, home resale is not allowed in this paper. Parent and adult children make the optimal consumption and saving decisions independently, and intergenerational transfer is realized through lump-sum fashion at each period. To quantify the effects of housing price and financial constraint on household life-cycle behavior, I calibrate the model using China Household Finance Survey data from 2011 and China Health and Retirement Longitudinal Study (CHARLS) survey from 2013. With the calibrated parameters, I then conduct three policy experiments to isolate the effects of decreasing housing prices and down-payment ratio on savings rate and household living arrangement patterns.

There are three main results in this paper. The first is that when down payment ratio drops from 30% to 20%, household behavior patterns on savings and coresidence change very little. With the house price staying the same, lowering down payment alone does not fundamentally change housing affordability level, hence the optimal timing for young adult to move out would not change very much. It is for the same reason, the life cycle savings profile for coresidence households stays roughly the same. The down payment does have greater effects on the savings profile of non-coresidence households during early periods, which is mainly caused by the fact that less initial investment is made due to the dropped down payment requirement. This result implicates that down payment has very limited effects on household savings rate and coresidence pattern when house price is very high. Therefore, the policies designed to derease down payment requirement will not solve the housing affordability problems in China.

The second result is regarding the impact of housing price on households age-savings pattern. When house price decreases by half, the savings rate of coresidence households decreases by roughly 27%, and the savings rate of non-coresidence household decreases by about 32%. Aside from the significant effects of house price on aggregate savings rate, housing has deeper impact on younger households in terms of savings behavior. There are three reasons responsible for the decline in household savings rate. The first is that cheaper housing requires less transfer from parents to adult children, thus parents are less obligated to save for intergenerational transfer. Secondly, adult children feel less compelled to reduce consumption in facilitating savings due to the mitigated credit constraint, permitting lower saving rate. Third, lowering housing price lead to less housing asset, thus the savings rate of non-coresidence households would drop over life cycle. On further examination, when house price and down payment decrease at the same time, the savings rate over life cycle drops even more. This result implies that people are strongly motivated to save for housing, hence the analysis of savings

54

cannot ignore housing choice. Meanwhile, in a life cycle model with endogenous housing choice, intergenerational transfer and coresidence are important elements that connect two generations.

The last result is that when housing price decreases by half and down payment drops from 30% to 20%, coresidence rate over life cycle would decline, but still remains higher than the U.S. level. I choose this policy experiment because it makes the housing market condition similar to that of the U.S., meaning households facing a house value-to-income ratio of 3 and 20% minimum down-payment[2]. The less-than-expected effect on coresidence probably has two reasons. First, this model does not include repeated home sales, hence people are tempted to buy houses later so that they can enjoy bigger or better houses. Second, since rental housing is not built in the model as an endogenous housing choice, young adults would have no choice but to share parent's house until they can afford homes. If the model is expanded to include both repeated home sales and rental housing, the coresidence pattern might be explained more.

In summary, the above results support the main hypothesis in this paper. Housing indeed is one of the most important motives for household saving, and it has more impact on young household's savings behavior than the middle-aged. Meanwhile, comparing to the insanely high house price, down payment has less influence on household's saving and living arrangement choices, which is consistent with literature. Lastly, the high coresidence rates in China are not solely determined by the housing shortage situation.

---

[2]A typical measure for housing affordability is house price-to-income ratio, which is dividing the median home price by median household income. And the mortgage market usually assumes a 30-year fixed rate mortgage with a 20 percent down payment, see http://www.freddiemac.com/finance/report/20160531_how_to_worry_about_house_prices.html.

Institutional Background and Literature

Before 1998, housing was allocated to the employees as an in-kind welfare of the public sectors and SOEs. After 1998, the employer-based public welfare housing system was terminated, and the occupied houses were sold to workers with heavily-discounted prices, as a result, more than 80% of the households own their houses. The privatization reform of housing market has led to a decade-long housing market boom and enormous housing price appreciation. According to Fang, Gu, Xiong, and Zhou (2016), housing prices in first tier cities like Beijing and Shanghai have risen by five folds from 2003 to 2013, and roughy three folds in second and third tier cities. On top of housing price appreciation, the down payment for housing mortgage loans is usually as high as, or even higher than 40%. The price-to-income ratio is normally used to measure the housing affordability level. With the combination of high housing price and high down payment, the price-to-income ratio is about 6 or higher for the middle income households, which indicates very severe housing affordability problems in China, suggesting that people need to save a lot to buy houses.

This paper is related to two branches of literature: the literature about the Chinese households saving behavior (e.g., Wei and Zhang (2011); Chamon and Prasad (2010); Choukhmane, Coeurdacier, and Jin (2013)) , and the literature on coresidence (Rosenzweig and Zhang (2014); Kaplan (2012)). Chamon and Prasad (2010) documented household saving behavior relative to age, time and cohorts. They found that the increased expenditure on education, health and housing has statistically significant effects on the rising savings rate. In addition, they suggested that credit constraint should be important in explaining the high saving rate of the young, but they did not discuss thoroughly the link between credit constraint caused by housing and household saving behavior. Wei and Zhang (2011)

argued that gender imbalance caused male individuals to save more so that they can improve their social status on marriage market. They believed this competitive saving motive has pushed the housing price in China to appreciate. However, the hypothesis of this paper is that housing is an important motive for household saving, and the main focus is to analyze the effects of housing price appreciation on household life cycle savings profile. Choukhmane, Coeurdacier, and Jin (2013) analyzed the impact of "one child policy" on household saving behavior. They believed that policy restriction on fertility caused education investment on child to increase due to the quantity substituting for quality effect, and the middle-aged households save more in anticipation of the reduced transfers from their only child. As a result, they estimated a "hump shaped" age saving pattern, which is drastically different from the predicted age-savings pattern in this paper. Another popular idea for the rising savings rate is changes in demographic, according to Curtis, Lugauer, and Mark (2015), the rising share of working population for the last decade can explain about half of the increase in savings rate. As for the relative literature on coresidence, the paper by Rosenzweig and Zhang (2014) is the only one I am aware of that studies the effect of intergenerational coresidence on China's particular age saving profile. They developed a multi-generation life-cycle model including both coresidence and saving, and tested the effects of coresidence on young's saving empirically. The difference between their paper and this paper is that they did not take into consideration of the effects of credit constraint.

The rest of the paper is organized as follows. Section 2.2 first gives some brief empirical analysis on the relationship between coresidence and saving behavior, and then introduces detailed facts on urban China's households' savings behavior, coresidence, housing affordability problem, and intergenerational transfer using CHFS and CHARLS datasets. Section 2.3 introduces the life-cycle dynasty model of endogenous housing choices, and some analytic implications of credit constraint on housing and savings decisions. Section 2.4 calibrates the model with CHFS and

CHARLS data. Section 2.5 conducts three policy experiments using the calibrated parameters. Section 2.6 presents an alternative approach for model calibration. Section 2.7 provides conclusion of this paper.

2.2 Data Introduction And Empirical Analysis

The purpose of this section is to introduce the main information about household savings, intergenerational transfer, coresidence and housing affordability problem in China. The China Household Finance Survey (CHFS) data from 2012 is a national representative sample that contains 8,438 households (Urban=5,194; Rural=3,244) and 29,500 individual observations, with detailed information on household asset, debt, income and expenditure,and demographics. Because CHFS lacks adequate information on intergenerational transfer, especially the type of transfer for aiding children's housing purchase, I use CHARLS (China Health and Retirement Longitudinal Study) to account for the extent of intergenerational transfer in urban China. From here on out, I only report the statistics on urban sample in CHFS and CHARLS. I do not include rural sample in my study, because things are quite different in rural area. Chinese rural households generally have more than one child, indicating that fertility choice is endogenous, which will complicate the model regarding intergenerational transfer and coresidence. Because parents need to decide how many children to have, and whether or not to live with their adult children, as well as which child to live with. Therefore the circumstances of adult children core siding with their parents in rural area is quite different from that of urban area.

One of the major pattern I could obtain from CHFS is age-savings profile. The data contains household level aggregate income and consumption expenditure, so I can get the saving rate (defined as $1 - \frac{consumption\ expenditure}{disposable\ income}$) at household level [3]. In

_____

[3]Consumption expenditure includes household expenditure on food, clothing, housing ex-

the case of coresidence, the aggregate multigenerational savings rate is available, but individual level savings rate is not included in the data. Besides saving, detailed information on age, coresidence and employment status for each family member is also available, so I could keep track of the age-coresidence pattern for each household[4]. Figure 1 shows the age-savings pattern using CHFS from 2012, by which we can see that the savings rate of the young (before age 40) are roughly the same or slightly higher than the middle-aged people (40 - 50).

The life cycle coresidence rate obtained from CHFS data is defined as the mean of probability that adult children living with their parents, where "parents" includes parents, parents-in-law, and grand-parents[5]. In addition, the calculation excluded the respondents who are students. We can observe a striking pattern of the age-coresidence profile from figure 2.2 : older people are more likely to live apart from their parents. Besides age, there are other factors affecting coresidence, such as income (richer people are less subject to credit constraint posed by housing) and geographic location (housing prices vary a lot across different regions), hence it is necessary to examine further the relationship between coresidence and age when household income and geographic region are taken into consideration. The coresidence rate is considered as a function of household's log-income and individual's age group dummies, as well as individual's economic region dummies. I regress the coresidence rates by the following regression for each individual who aged between 20 to 45 in urban China:

---

pense (housing renovation or maintenance), family equipment (including consumer durables) and daily necessities, home services (e.g., cleaning, cooking and child-sitting), medical expenses, communication, travel expense, recreation, education and childcare.

[4]The labor income of family members usually just report household head or spouse's labor income, so I could not identity children and parent's labor income at the same time for most of the households.

[5]This definition is corresponding to the obtained pattern in figure 2.

Table 2.1: **OLS regression for coresidence**

Dependent variable : Indicator for coresidence	
Age group 1 (20~23)	0.58***
	(0.039)
Age group 2 (24~26)	0.49***
	(0.039)
Age group 3 (27~29)	0.50***
	(0.04)
Age group 4 (30~32)	0.39***
	(0.039)
Age group 5 (33~35)	0.26***
	(0.043)
Age group 6 (36~38)	0.21***
	(0.041)
Age group 7 (39~42)	0.09**
	(0.039)
Log household income	-0.02***
	(0.008)
House_value (Current)	0.02***
	(0.006)
Intercept	0.27
	(0.102)
City Fixed Effect	Y
R^2	0.1663
Observations	2857

$$C_i = \beta_0 + \sum_{s=1}^{8} \beta_s \cdot a_s + \sum_{r=1}^{5} \alpha_r \cdot d_r + \gamma \log(income_i) + \epsilon_i;$$

where C_i is an indicator variable that takes a value of 1 if living with parents, a_s is an indicator variable that takes a value of 1 if the individual belongs to age group s. There are 8 age groups: age 21-23, age 24 - 26, age 27 - 29, age 30 - 32, age 33 - 35, age 36 - 38, age 39 - 42, age 43 - 45. d_r is an indicator variable that takes a value of 1 if the individual lives at region r, and the 5 economic regions are: mega-city (Beijing, Shanghai, Guangzhou, Shenzhen); east coast; central China; northeast China; and western China. The reason to include economic region in the regression is that there is a wide variation in housing prices across different regions,

since high housing prices is one of the causes for people choosing coresidence, disregarding economic regions might lead to biased estimation. From table 2.1, we can see that the probability of one living with one's parents is decreasing in age, and people tend to live apart in richer households. In addition, table 2.1 suggests that young adults tend to move out earlier in richer households; but they tend to stay longer if the house has greater value. Based on the regression results in table 2.1, I can obtain the predicted age-coresidence pattern when controlling for other factors, which is shown in figure 2.3. According to figure 2.3, we can see that the life-cycle coresidence pattern still remains after controlling for household's income and economic region.

Figure 2.3: **Life-cycle pattern of coresidence**

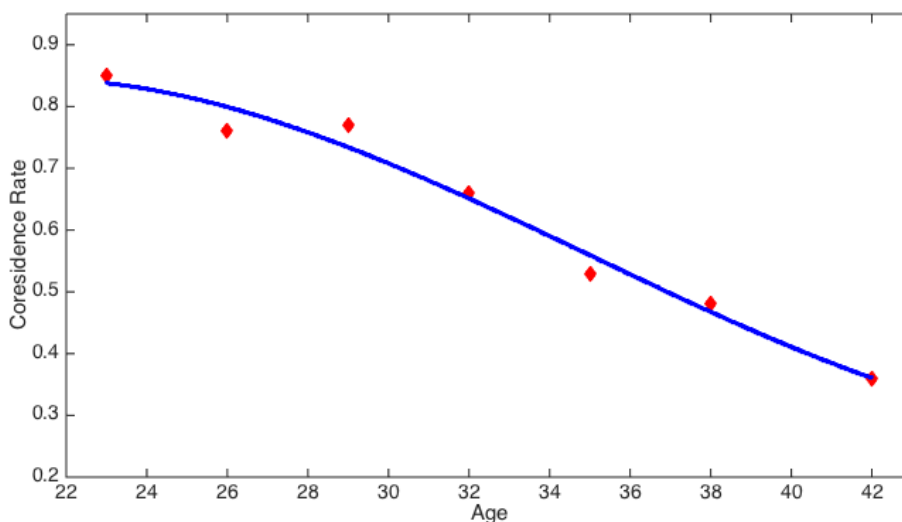Figure 2.4 shows the age-savings profiles under different household structures: coresidence and non-coresidence. We can see that the savings rate of the young under coresidence is higher than that of non-coresidence households, and the savings rate has a down-ward trend as people get older. Because there are other confounding factors affecting household saving, such as household income, housing

61

Table 2.2: **Impact of Coresidence on Household Saving Rate**

Dependent variable: household Saving Rate	
Coresidence	0.23**
	(0.11)
Coresidence ×age	-0.005*
	(0.003)
Age	0.009***
	(0.002)
head_education	-0.87***
	(0.009)
Log household income	0.76***
	(0.018)
Log house value (Current)	-0.13***
	(0.01)
Observations	3353
Adjusted R^2	0.3657

asset and age, it is necessary to examine whether the above described pattern still exists with empirical analysis. With the data on household savings rate and coresidence, I regress savings rate on coresidence indicator variable with the following equation :

$$saving_i = \beta_0 + \beta_1 coresidence_i + \beta_2 coresidence_i * age_i + \beta_3 \log(income_i)$$
$$+\beta_4 age_i + \beta_5 education_i + \beta_6 house_i + \epsilon_i$$

Here $coresidence_i$ is an indicator variable, with $coresidence_i = 1$ stands for adults children and parents living together, and $coresidence_i = 0$ stands for living apart. In addition, $house_i$ represents log house-value, and $education_i$ stands for household head's education level. Table 2.2 shows that the coefficient of coresidence is positive, and the coefficient on the interaction between coresidence and age is negative, which means coresidence and household savings rate are positively correlated when adult children are young. And as young adults age, this positive correlation weakens.

Figure 2.4: **Age-savings profile under coresidence and non-coresidence**
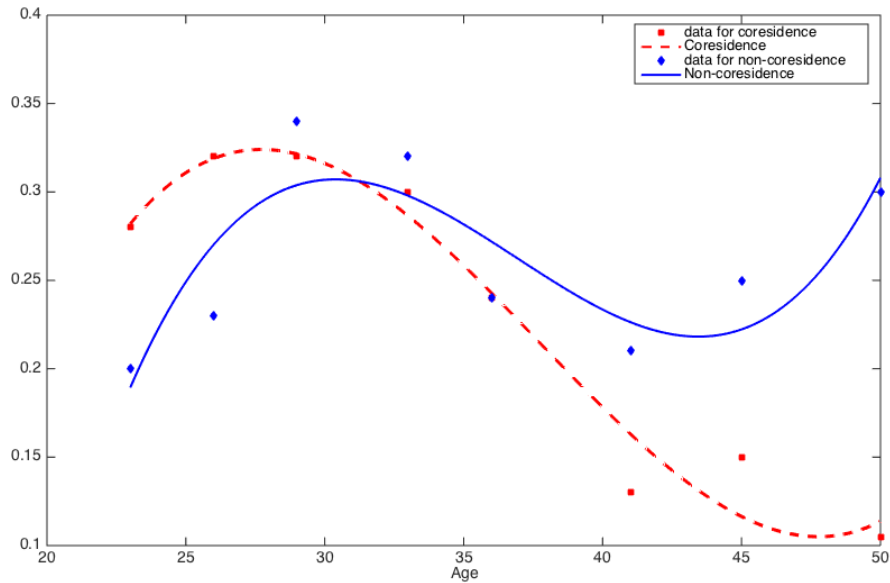
China Health and Retirement Longitudinal Study (CHARLS) is a biennial survey that is nationally representative of residents age 45 and older. There are two national baseline survey: 2011 and 2013, only 2013 survey contains relevant data on intergenerational transfer from parents to adult in supporting children's housing purchase. In China, the financial support adult children get from their parents when they buy houses is substantial and quite common. However, the survey has no exact information on the level of transfer just for housing support, such that households were not asked whether they bought their own housing, or accepting the housing as gifts from their parents, or how much financial help they received from their parents. Hence I use wedding gifts as an indicator for intergenerational transfer in this paper. In the event of wedding, Chinese parents usually give their children a large amount of money transfer as betrothal gifts. Some parents even buy a house for their children. This common practice of wedding transfer gives indication as to why the young would want to save or buy a house. One important reason is marital competition, as suggested by Wei and Zhang (2011). Because

63

Table 2.3: **Intergenerational Transfer Measured by Wedding gift**

Family Income (Quantile)	Share of family gave gift to children	Ratio of gift value to parent's income	Share of family buy house for children
1(Low)	72%	3.7	10%
2	70%	2.4	11%
3	75%	1.5	15%
4(High)	77%	1.9	17%

Note: table 3 is derived using CHARLS in 2013 , and only including urban households whose children's wedding occurred after year 1998.

owning a house will significantly improve one's status on marriage market, thus savings rate is elevated in households with young adults. The 2013 CHARLS data asked specific questions on intergenerational transfer for wedding, such as "Did you give betrothal gifts when your child got married", "How much was the total value of those betrothal gifts", "Did you buy a house for him/her when your child got married" and "How much was the total value of the house". Table 2.3 presented the extent of intergenerational transfer using wedding gifts measurement [6]. We can see from table 2.3 that the intergenerational transfer from parents to children is remarkably high.

2.3 Model

To examine the relationship between housing prices and savings in a regime with intergenerational transfer and coresidence, I construct a simple life-cycle dynastic model. Parents and adult children separately determine their optimal consumption paths and whether to share the parents' houses. Parents have houses to begin with, so they can provide housing support to adult children in the form of coresidence. Meanwhile, they can make intergenerational transfer in aid of children's housing purchase. The option of coresidence allows adult children to

---

[6]Some households have more than one child, in that case I use the maximum of wedding gifts that parents gave all the children given that the wedding occurred after 1998.

save living costs through shared housing, and delay the timing of housing purchase. Eventually adult children will have enough savings, at least enough to pay for the down payment, then they will move out of parent's house. There are two living arrangements for adult children: either living with their parents (coresidence) or buying their own houses. This set-up categorizes households in two groups, coresidence and non-coresidence, which allows me to explore the effects of house prices on both groups. Meanwhile, I can get implications of house prices and borrowing constraint on China's coresidence pattern.

To highlight the multiple links between house prices, savings and coresidence, I make a number of simplifying assumptions. First, in consistent with the one-child policy in China, each household has only one child and no fertility choice is involved. Second, repeated home sales and housing rental market are not included in this model[7]. Even though these two features on housing market may seem unusual, there are several other papers adopt the same assumptions, such as Fang et al. (2016). The model has two generations: parent (p) and children (k). The timing is simple: at the beginning of period t, if adult child has not bought his own house yet, he must decide between coresidence and buying. If coresidence, adult child can live with his parent freely at period t, then he needs to decide again between coresidence and buying at next period $t + 1$; if instead he decides to buy his own house at period t, then he lives in that house until the last period T. To make the model closer to reality, adult child has two choices regarding his own house at last period T: he can leave his house to the grandchild as a bequest, or he can sell the house and consume everything.

The life-time utility function of household i is given by:

---

[7]In urban China, about 15% households rent their houses, among which two thirds are public rental housing (public rental housing is a form of subsidized rental housing provided by government for the low-income people), therefore private rental housing represents a very small share on Chinese housing market. On the other hand, due to the nascent nature of the Chinese housing market, only 2% of the total housing stock in Urban China is accounted by repeated home sales, thus repeated home sales is not included in the model.

$$V_i = \sum_{t=1}^{7} \beta^{t-1} U\left(c_{i,p}^t, h_{i,p}^t\right) + \sum_{t=1}^{T=12} \beta^{t-1} \eta U\left(c_{i,k}^t, h_{i,k}^t\right);$$

where β is the discount factor, η is the altruism parameter of parent. I assume utility function takes the form of $U(c,h) = \gamma \frac{c^\delta}{\delta} + (1-\gamma)\frac{h^\delta}{\delta}$, where c is non-housing consumption and h is housing consumption[8]. Household in the model is indexed by i, and identified by four initial parameters: adult child's initial income $y_{i,k}^1$, old parent's initial income $y_{i,p}^1$, old parent's initial asset $s_{i,p}^0$, and old parent's house $\bar{h}_i$. Parent's house $\bar{h}_i$ is fixed for each household i and positively correlated with parent's income. At period 1, income profile $(y_{i,k}^1, y_{i,p}^1, s_{i,p}^0)$ is realized, which is drawn from a multivariate normal distribution, with the correlation coefficient between child and parent income ρ and correlation coefficient between parent's income and parent asset ρ^{pa}. Parents live for 7 periods, with each period equivalent to 5 years; they start at age 45, retire at age 60, and die at age 80 with certainty. Adult children live for 12 periods (T=12) : start from age 20, also die at age 80. Since parent/children living together implies inconvenience or lack of privacy for both, there is a disutility factor λ in the case of coresidence: $h_{i,p}^t = h_{i,k}^t = \frac{\bar{h}_i}{\lambda}$, $\lambda > 1$. With the above assumptions, the problem of owning housing at period t is defined as following:

$$V_i = \max_{\left\{c_{i,k}^t, s_{i,k}^t, d_i^t, h_i^t\right\}_{t=1}^{T}, \left\{c_{i,p}^t, s_i^t\right\}_{t=1,}^{7}, G_i, \psi_{i,k}^T} \sum_{t=1}^{7} \beta^{t-1} \left[\gamma \frac{\left(c_{i,p}^t\right)^\delta}{\delta} + (1-\gamma)\frac{\left(h_{i,p}^t\right)^\delta}{\delta}\right]$$

$$+ \sum_{t=1}^{12} \beta^{t-1} \eta \left[\gamma \frac{\left(c_{i,k}^t\right)^\delta}{\delta} + (1-\gamma)\frac{\left(h_{i,k}^t\right)^\delta}{\delta} + \psi_{i,k}^T \kappa \frac{\left(h_{i,k}^t\right)^\delta}{\delta}\right];$$

---

[8]Unlike most literature on housing demand, housing in this paper is measured by housing area rather than housing value.

where

$$h_{i,p}^t = \begin{cases} \bar{h}_i & if\ non-coresidence \\ \\ \bar{h}_i/\lambda & if\ coresidence \end{cases} \quad ; h_t^{i,k} = \begin{cases} h_{i,k} & if\ non-coresidence \\ \\ \bar{h}_i/\lambda & if\ coresidence \end{cases} .$$

Here κ is the altruistic parameter for housing bequest, and $\psi_{i,k}^T$ is indicator for the option of leaving bequest, with $\psi_{i,k}^T = 1$ representing adult children choosing to leave his purchased house $h_{i,k}$ for grandchildren, whereas $\psi_{i,k}^T = 0$ representing adult child selling the house rather than leaving it as a bequest for grandchild at last period T. The first decision adult child needs to make is choosing the optimal timing of housing purchase, which is d_i^t, with $d_i^t = 0$ representing not buying and $d_i^t = 1$ representing buying. The timing of housing purchase must be before $t \leq 7$, since old parent will die after period 7. If adult child still has not owned his own house at period 7, he will inherit his parent's house $\bar{h}_i$ and not buying at all. Given the optimal timing of housing purchase, at period t, budget constraints are as following:

$$c_{i,p}^t + s_{i,p}^t = (1+r)s_{i,p}^{t-1} + y_{i,p}^t - G_i; \qquad t \le 7$$

$$c_{i,k}^t + s_{i,k}^t + d_i^t P^t h_{i,k} = (1+r)s_{i,k}^{t-1} + y_{i,k}^t + G_i; \qquad t \le 7$$

$$G_i \ge -\bar{G}; \qquad t \le 7$$

$$c_{i,k}^t + s_{i,k}^t = (1+r)s_{i,k}^t + y_{i,k}^t; \qquad t > 7$$

$$c_{i,k}^T + s_{i,k}^T = (1+r)s_{i,k}^{T-1} + y_{i,k}^T + P^T h_{i,k}; \qquad if \ \psi_{i,k}^T = 1$$

$$c_{i,k}^T + s_{i,k}^T = (1+r)s_{i,k}^{T-1} + y_{i,k}^T; \qquad if \ \psi_{i,k}^T = 0$$

$$s_{i,k}^t \le -(1-\alpha)P^t h_{i,k}; \qquad if \ d_i^t = 1$$

$$P^t = (1+g)P^{t-1}; \qquad t \le 7$$

$$P^t = (1+2\%)P^{t-1}; \qquad t > 7$$

Income grows with a fixed rate g, hence adult child's income from period 1 to period 8 is given by $y_{i,k}^{t+1} = (1+g)\, y_{i,k}^t$ for $2 < t \le 8$; after period 8, adult child's retirement income is given by $y_{i,k}^t = \vartheta y_{i,k}^8$ for $9 \le t \le T$, where ϑ representing replacement ratio. Parent's income is such that: $y_{i,p}^{t+1} = (1+g)y_{i,p}^t$ for $t \le 3$, and $y_{i,p}^t = \vartheta y_{i,p}^3$ for $4 \le t \le 7$. In the budget constraint, α is minimum down-payment ratio, P^t is the fixed housing price at period t, which grows with the same rate as household income g until period 7, then its growth rate drops to the steady state level 2%[9]. G_i stands for intergenerational transfer, and it is the same amount each period as long as $t \le 7$. Note that the intergenerational transfer G_i can be both positive (with positive meaning transfer is given from parent to adult child) and negative (meaning transfer from adult children to parent), but it has a lower limit $\bar{G}$, indicating transfer made from adult children to parent can be positive but with limited amount. This assumption coincides with the fact that intergenerational transfer in China from adult children to older parents are usually limited in value,

---

[9]There are two approaches to deal with housing price progress in this paper. The first is assuming house price grows in same rate as household income before $t \le 7$, and the growth rate drops to 2% after $t > 7$; in which case the degree of housing affordability is constant for household because adult children can only buy houses before period 7. The second approach is assuming housing price grows with the same rate (8%) of household income through adult children's entire life cycle. I will present the results of first approach in the main paper, and the results of second approach can be found in section 6.

however it can be a huge amount the other way around. In this model, household's asset includes both financial asset and housing asset. Thus adult children's savings at period t is defined as:

$$
\begin{cases}
save_{i,k}^t = s_{i,k}^t - s_{i,k}^{t-1} & if \ d_t^i = 0 \\
save_{i,k}^t = s_{i,k}^t - s_{i,k}^{t-1} + P^t h_{i,k} & if \ d_t^i = 1
\end{cases}
.
$$

2.4 Model Calibration

In this section, I discuss the choice of the parameters I used in the model, and the calibration procedure. I then show the life-cycle profiles generated by the calibrated parameters. The model parameters are calibrated to the life cycle profiles on coresidence, housing expenditure, savings rate and intergenerational transfer using data from CHFS and CHARLS. I then conduct three policy experiments to discuss the implications of changing down payment and housing prices on household saving and coresidence patterns.

2.4.1 Baseline Parameter

Now I introduce my choice of parameters, shown in table 2.4. Some parameters I take directly from the pre-existing literature or from the data, and the remaining parameters need to be calibrated using the data. Annual real household income growth rate is taken to be 8%, which is calculated from survey data China Family Panel Studies (CFPS) [10]. The annual interest rate 4%, same as in Curtis, Lugauer

---

[10]The real household income growth rate is taken to be 8%, the same growth rate of real housing price in China according to Fang (2015), since house prices and household income have similar trend over the last decade. Plus, according to Song (2010), a 8% income growth rate is within reasonable range. In this paper, household income contains both labor income and other sources of income, such as transfer and capital return.

and Mark (2015). The minimum down payment ratio as 30%, according to the regulation in China's housing mortgage market. The pension replacement rate i set to be 35%, same as the replacement rate after adjusting for pension coverage in Song and Yang(2010). From CHFS survey, I can only observe the individual incomes of both adult children and parents when they live together, by calculating which, I get the correlation coefficient between income(log) of parents and adult children 0.217[11], and I take this as the simulated target to get the correlation coefficient between parent and adult children's income. The mean (μ^k, μ^p, μ^a) and standard deviation $(\sigma^k, \sigma^p, \sigma^a)$ of initial log-income of parent, log-income of adult children and initial log-asset of parent are taken directly from CHFS urban data. The correlation coefficient between parent's (log) income and asset (ρ^{pa}) is taken from CHARLS urban sample, only including those whose children have not been married yet at the survey year, so that the asset data can reflect the savings parents have accumulated for the event of their children's wedding.

With the means and standard deviations for children's initial income as well as parents' initial income and asset, I draw a multivariate normal distribution to pin down the initial income distribution among households, then given the fixed income growth process and the rule of retirement income, I thus have the life-time income profile for all households. The price index is calculated in the following approach : $P^1 = \frac{\Gamma \bar{y}_k}{h}$, where $\Gamma = 6$ stands for the average price-to-income ratio, $\bar{y}_k$ stands for the mean of child's initial income, and $\bar{h}$ denotes the mean housing quantity measured from the data, which is 10 units, equivalent to 100 square meters. The calibrated housing price index P^1 is not real housing prices, just indexed prices to keep households facing certain level of price-to-income ratio.

---

[11]Using CHARLS data, individual incomes of both child and parent are reported in the coresidence households, hence I can get the correlation coefficient for these households, but if adult child and parents live apart, there is no way to determine their individual incomes. The correlation between log-income of parent and children ρ is treated as free parameter so that it fits the correlation coefficient under coresidence in the calibration.

Table 2.4: **Exogenous Parameters**

(Parameters taken from the previous literature or data)		
Length of periods	T	12
Annual interest rate	r	4%
Down payment ratio	α	30%
Annual income growth rate	g	8%
Pension replacement rate	ϑ	35%
Mean of initial log-income for child (age 20 - 30)	μ^k	10.77
Mean of initial log-income for parent (age 40 - 50)	μ^p	10.7
Mean of initial log-asset for parent	μ^a	12.46
Standard deviation of log-income for child	σ^k	1.18
Standard deviation of log-income for parent	σ^p	1.12
Standard deviation of log-asset for parent	σ^a	1.54
Correlation coefficient between parent's (log) income and asset	ρ^{pa}	0.327
Housing price index at initial period	P^1	4
($P^1 = 5$ converted to dollar is about 60 $ per square foot)		

In the model, coresidence decision depends on parent's initial housing condition, because adult children are more likely to live with parents if their parents have bigger or better houses. CHFS data asks detailed questions about housing condition, such as the floor area, the number of bedrooms, and the original cost of housing. Due to the difficulty of measuring housing quality, I use housing floor area as the indicator for housing quantity, the variation in housing area is not very large, such that the mean of housing area is 100 m^2, and only about 5% households live in houses larger than $200m^2$ [12].

2.4.2 Calibrated Parameters and Predicted Life Cycle Profiles

Given the parameters in Table 2.4 and the life-cycle income process described in section 2.4.1, I can simulate the model and generate life-cycle profiles for coresidence, housing expenditure, savings rate, and intergenerational transfer. The

---

[12]I categorize the households in four income quartiles, the mean of housing area within each quantile (from low income to high income quantile) are 112, 118, 125, 130 square meters respectively.

calibration aims to fit 33 moments from the data: the mean of coresidence rate for each period (there are 7 periods in total), the mean of housing value-to-income ratio for each period, the life cycle savings rate for both coresidence and non-coresidence households, the ratio of average intergenerational transfer over parent's initial income for 4 income quantiles, and the correlation coefficient between log-income of parent and children under coresidence.

There are 7 parameters that needed to be calibrated: the time discount factor β, disutility factor of coresidence λ, utility coefficient of consumption γ, the measure for elasticity of substitution δ, parent's altruistic parameter η, bequest parameter ϑ and the log-income correlation between parents and children under coresidence. Table 2.5 reports the parameters that I calibrated by matching certain moments from data.

Figure 2.5: **Model-Predicted and Actual Life Cycle Patterns**

Table 2.6: **Estimated Parameters**

Discount factor	β	0.99
Relative importance of non-housing	γ	0.75
Elasticity of substitution between housing and consumption	$\frac{1}{1-\delta}$	0.55
Parent's altruism	η	0.80
Disutility of coresidence	λ	2.43
Bequest parameter	ϑ	1.43
Income correlation between children and parent	ρ	0.37

This exercise is to check the performance of the calibration model by showing the contrast of simulated results and data moments. Also it is useful to provide some intuition for the theoretical model and relevant policy experiments. Figure 2.5 presents the simulation results on life-cycle profiles : life-cycle coresidence pattern (upper left plot), life-cycle profile on housing expenditure that measured by house value-to-income ratio (upper right plot), life-cycle savings rate profile for coresidence household (middle left), life-cycle savings rate profile for non-coresidence household (middle right), and the intergenerational transfer conditional on parent's income quartile (bottom right plot). The model is able to generate life cycle patterns quite comparable to the data, except for a few discrepancies. The model-generated coresidence pattern has a downward trend over time, but with a slower convergence rate. The model-generated house value-to-income ratio has a similar trend to the data over life cycle, but it is slightly higher for the middle-aged households. The model-generated age-savings profile for non-coresidence household has lower savings rates at later stages of life. The model-generated intergenerational transfer is lower than the data for the low and middle income households. In general, the fit between model and data is quite close, so this model is suitable to analyze the implications of credit constraint on life-cycle profiles.

73

2.5 Policy Experiments

One of the goals of this paper is to separate the effects of housing price and financial constraint (down payment) on savings rate as well as household structure, so that the possible channel of housing condition affecting household saving would be identified. In this section, I conduct three counterfactual policies. Policy one: lower the downpayment ratio from 30% to 20% and keep other parameters the same; Policy two: lower the housing prices by half and keeping down payment and other parameters the same time; Policy three: lower price by half and drop down-payment ratio to 20%, so that households face house price-to-income level and down payment ratio similar to that of the United States. For each policy I simulate a set of predicted moments of life-cycle profiles, and then compare the results to those of the baseline moments [13].

### 2.5.1 Policy One: Lower Downpayment Ratio from 30% to 20%

In baseline model, there are two sources causing housing affordability issue: high downpayment ratio and high housing prices. This policy focus on the effects of downpayment ratio only. When down payment is dropped from 30% to 20% [14], by fig. 2.6, coresidence rate over life cycle barely dropped at all. In this case, housing price is still extremely high, lowering down-payment alone is not enough to mitigate the financial burden, thus young adult would not change the optimal timing of housing purchase very drastically. Another interesting point captured from this experiment is that down-payment ratio decrease has very little effect on savings rate of coresidence household, but it has bigger impact on non-coresidence

---

[13]I use the same household income draws in the baseline and policy simulations.

[14]Here setting 20% downpayment ratio aims to implicitly compare the housing market of China and United States.

household especially in early period, see fig 2.6. The greater impact of down payment on non-coresidence household is not surprising, because now the share of housing asset accumulated in early stage is decreased due to the decreased minimum down payment requirement. This indicates that when housing price is very high (the house value-to-income ratio is about 6 in this case), policies designed to cut down payment would probably have limited impact on the housing affordability issue.

Figure 2.6: **Policy One: Down Payment Drops from 30% to 20%**
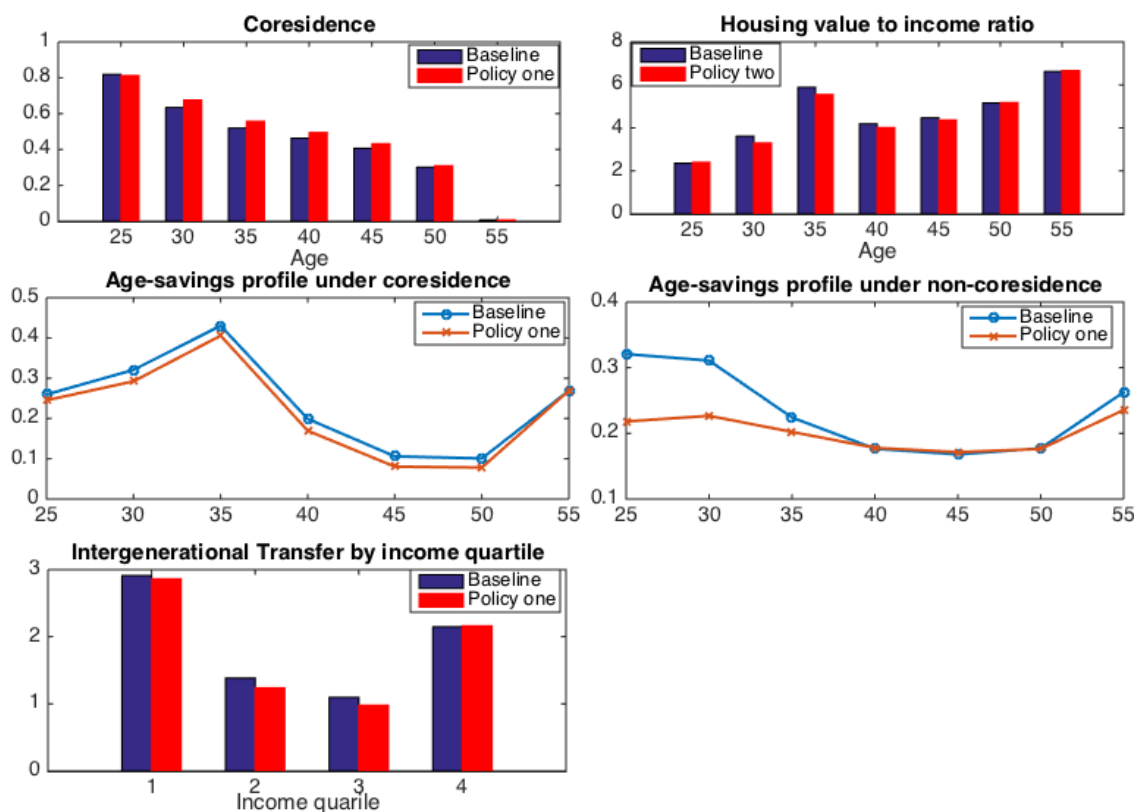

2.5.2 Policy Two: Lower House Price by Half

In this counterfactual experiment, down-payment remains 30%, and the housing

price decreases by half, which results the house value-to-income ratio dropping to 3. As you can see, the decreasing housing price has significant impact on the savings rate of both coresidence household and non-coresidence household over life cycle, and the effect is greater when household is at early stages of life cycle. On average, the savings rate of coresidence household dropped by 27%, and the savings rate of non-coresidence household dropped by 32%. In this experiment, cheaper housing indicates that parents need to save less for transfer purpose, which is verified in fig 2.7 (bottom right figure of intergenerational transfer). On the other hand, adult children are less financially constrained, they don't have to reduce consumption severely to ease the pressure of housing purchase. Moreover, because housing asset is included in household saving by definition, hence lowering housing price would lead to lower housing asset, then leading to lower saving of household. The lowering housing asset can be verified by the decreased house value-to-income ratio profile in fig. 2.7 (upper right figure).

### 2.5.3 Policy Three: Lower House Price by Half and Drop Downpayment Ratio from 30% to 20%

There have been popular concerns that if the housing bubble bursts in China, the Chinese economy will be severely damaged. Whether the rapid housing price appreciation represents housing bubble is not an issue in this paper, instead, this counterfactual analysis focus on the effects of housing price depreciation on the living arrangements as well as savings behavior at household level. In this policy experiment, I relax the credit constraint by decreasing the down payment ratio and housing price at the same time. The purpose of this counterfactual experiment is to simulate life-cycle profiles on savings rate and coresidence when housing market

Figure 2.7: **Policy Two: Housing Price Decreases by Half**
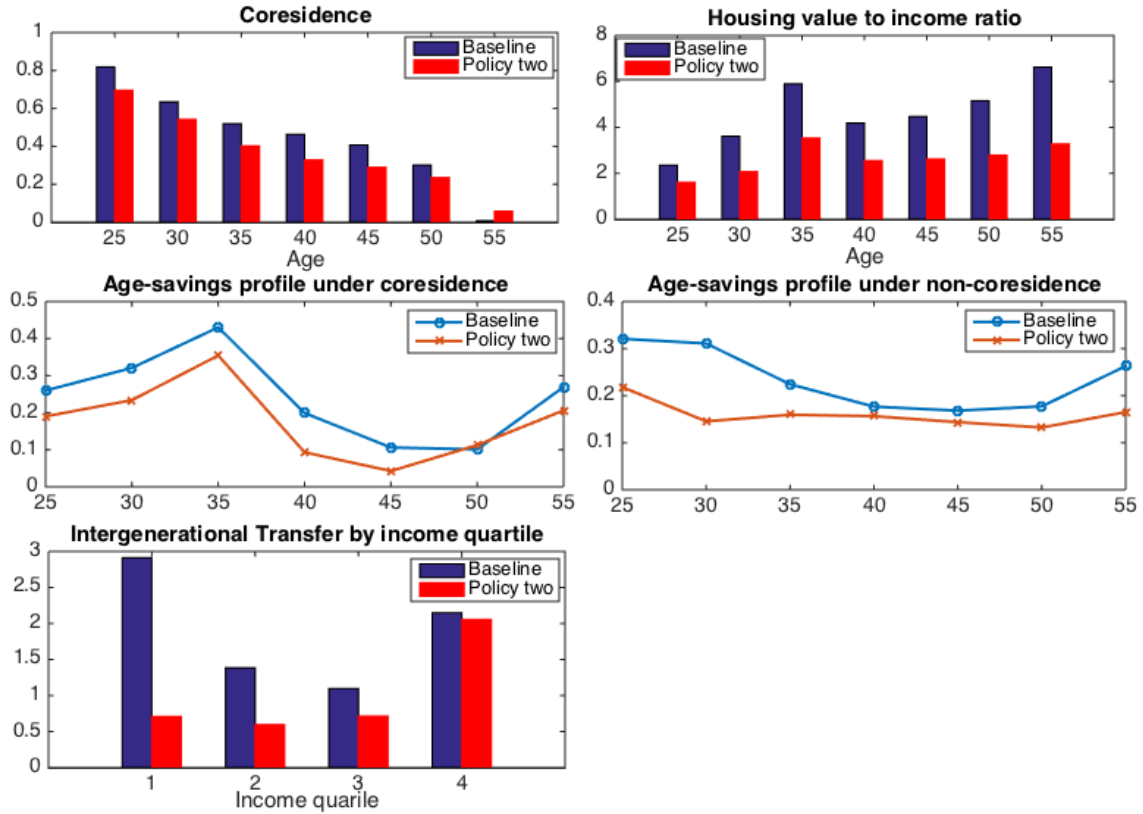
condition is similar to that of the U.S. housing market[15]. Figure 2.8 shows that the coresidence rates declined in all periods compared to baseline model. However, comparing to the life cycle coresidence profile of United States, see fig 2.9, the simulated coresidence in this case is still higher than that of U.S. over life cycle. There are several possible explanations for this situation. Because the model does not include repeated home sales, people probably want to buy houses later so that they can enjoy bigger or better houses. Secondly, since rental housing is not part of the endogenous housing choice in the model, young adults would have no choice but to live in parent's houses, until they save enough and buy homes themselves. If the model is expanded and includes home resales/renting, then the life cycle

Figure 2.8: **Policy Three: Housing Price Decreases by Half and Down Payment Drops to 20%**
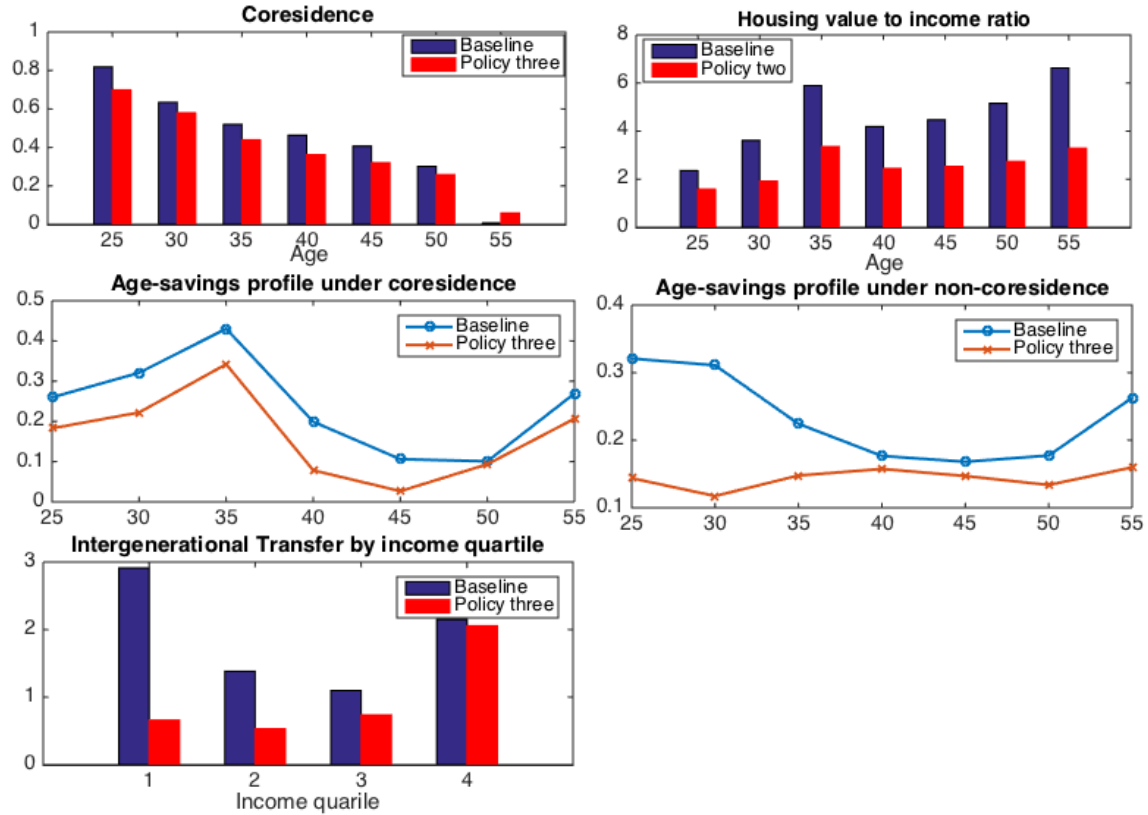

coresidence pattern might be fully explained.

Compared to counterfactual experiment 2, when both house price and down payment ratio are decreased simultaneously, the savings rates over life cycle decrease even more. On average, the savings rate of coresidence household dropped by 32%, and the savings rate of non-coresidence household dropped by 39%. The main take-away from this experiment is that the effect of housing price and financial constraint (caused by down-payment) on household savings rate is too big to be neglected. Therefore, we can say with confidence that housing is indeed one of the most important motives for saving.

---

[15]In U.S., the house price-to-income ratio for the middle-income group is about 3, and minimum down payment ratio is about 20%.

Figure 2.9: **Coresidence over Life Cycle When Housing Market in China is Comparable to U.S**
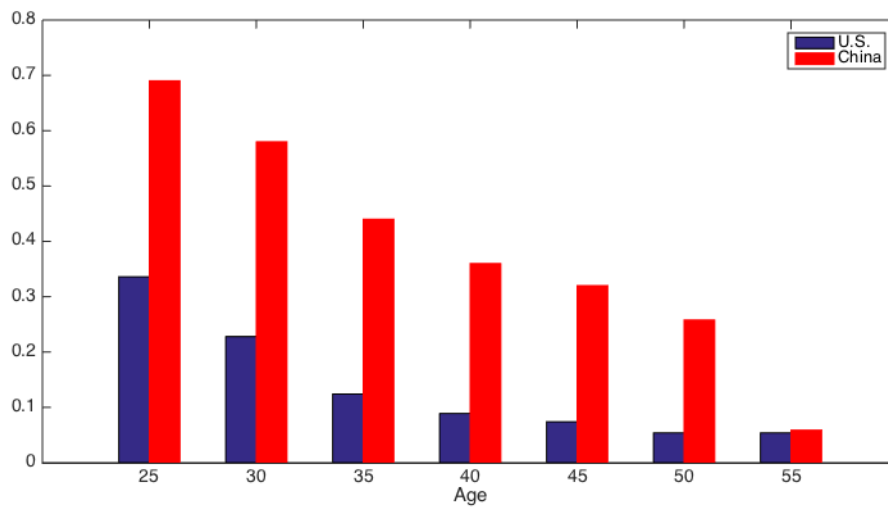
2.6 Alternative Calibration Using A Different Housing Price Appreciating Trend

Up to this point, the housing price appreciation is assumed to follow the same trend as the household income growth over life-cycle. With this assumption, the investment return from buying houses will not be blown out of all proportion, and households can behave rationally. Now I provide another approach for model calibration, which entertains the possibility of housing bubble. In this scenario, house prices increase at a steady rate 8%, and this trend will last forever, thus buying houses promises huge returns for households. This assumption will change age-savings pattern drastically, since capital returns usually decrease households savings. On the other hand, the coresidence pattern will also be different, because households now need to consider the tradeoff between credit constraint and future housing returns.

The procedure of calibration is very similar to that of section 2.4. I use the same parameters from table 2.4 to pin down the household characteristic parameters,

Table 2.8: **Estimated Parameters**

Discount factor	β	0.99
Relative importance of non-housing	γ	0.75
Elasticity of substitution between housing and consumption	$\frac{1}{1-\delta}$	0.54
Parent's altruism	η	0.89
Disutility of coresidence	λ	2.6
Bequest parameter	ϑ	1.25
Income correlation between children and parent	ρ	0.38

then simulate the model to fit exactly the same moments that depicting life cycle behavior patterns. The seven parameters that needed to be calibrated from the simulation are presented in table 2.6. The comparison between model simulated life cycle profiles and data is shown in figure 2.10.

There are three discrepancies suggesting the model does not fit the data quite well. Firstly, the model-simulated life cycle coresidence profile presents an uneven pattern. At the first period, the coresidence rate is extremely high, because only the richest households choose to buy houses. Next, there is a sudden big drop in coresidence rate at the second period, but after that the coresidence rate decreases slowly over time. In order to enjoy a higher housing asset return, people want to buy houses as early as possible, but at the same time they also want bigger houses since home resale is not allowed in the model. As a result, these two conflicting effects give rise to the simulated coresidence profile. The second significant discrepancy comes from the simulated age-savings profile for coresidence households, which is closely related to the coresidence pattern. In the beginning, only the rich young adults who are not financially constrained choose to buy houses, hence the savings rate is low; during later stages, the promised high capital returns motivate the lower-income/middle-income households save more, because they want to buy bigger houses, thus leading to an inverted "U" shaped pattern. Lastly, the model-simulated age-savings profile for non-coresidence housesholds is quite different from the data : it has a big spike at the second period, and then quickly

drops to a much lower level. The sudden jump in savings rate at second period is caused by the fact that about half households buy homes during that time, thus leading to a rapid increase of housing assets; after the second period, non-coresidence households save less due to the anticipation of future capital returns. Figure 10 suggests that a life cycle model with housing bubble does not fully or accurately capture the features of household's saving and housing behavior.

Figure 2.10: **Alternative Model-Predicted and Actual Life Cycle Patterns**
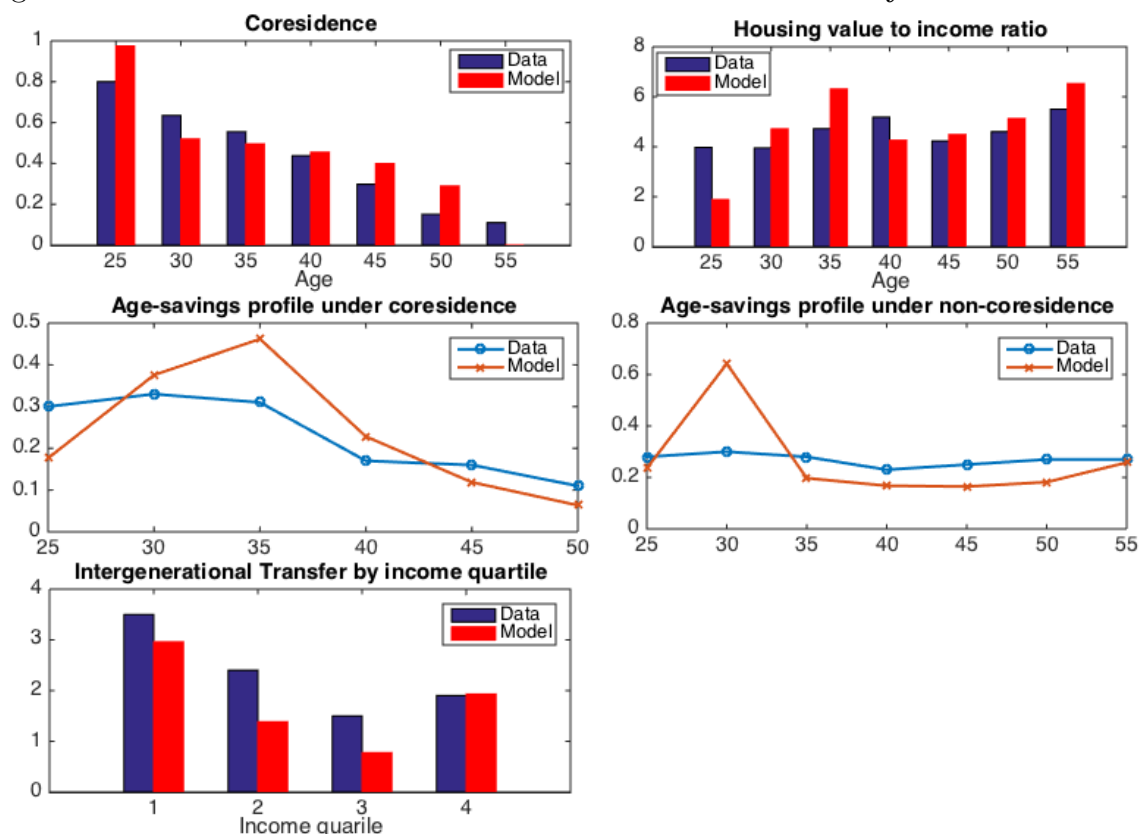
2.7 Conclusion

This paper concerns two interesting facts of Chinese urban household's living arrangement and savings behavior: the high savings rate (especially that of the young household), as well as the common situation of coresidence. Some papers

have tried to understand China's puzzling age-savings profile as the result of cohort, policy or culture effects. This paper quantifies the life cycle effects of housing savings motive on household savings rate by using a dynastic model which includes endogenous choices on coresidence and intergenerational transfer, credit constraint resulting from housing market condition.

There are three possible channels in which housing market condition might be important in explaining the high household savings rate in China. The first channel is through household living arrangement, more precisely, through coresidence. It is not uncommon for adult children to live with parents when financial burden of housing purchase is severe. Due to the consistent housing price appreciation in China, more than half of population has housing affordability problems. The down payment alone would require about two years of household income, thus coresidence provides a cheaper way of living, which indicates adult children can save more when living with parents. The positive correlation between coresidence and savings rate is examined in section 2.3, see table 2.3. The quantitative model provides a better understanding of the inter-relationship between coresidence and saving. According to the counter experiments in section 2.6.2 and 2.6.3, we can see that both coresidence and savings rate decreased as a result of decreasing housing price and down payment ratio, confirming the link between coresidence and high savings rate under the context of high housing price.

The second channel is through intergenerational transfer. Survey data of CHARLS (2013) gives indisputable evidence of intense transfer from parents to adult children for housing purchase motive (measured by wedding gift). When housing price is remarkably high, parents also need to save more in order to realize the heavy transfer, thus implying elevated savings rate among households. However, this transfer might happen for other reasons, such that altruism culture or tradition dictates that parents transfer a lot to children. But the simulated model supports

the hypothesis of transfer for housing purposes: the intergenerational transfer decreased greatly for all income-quartiles household when housing price decreases by half.

The last channel for housing motive affecting savings rate is through credit constraint: adult children are forced to save more so that they can pay for the down payment and mortgage. In the simulated experiment, when the minimum down payment ratio decreased from 30% to 20%, savings rate of household decreases under non-coresidence in early life cycle, but stays roughly the same under coresidence, which implies that down payment alone has limited influence over savings rate; However, when both down-payment and housing price are decreased, household savings rate decreases greatly over life cycle. This means that savings rate is indeed significantly affected by housing market condition, especially when credit constraint exists. This paper provides a simple life cycle framework to examine the influence of housing motive on household savings rate, which proves to be effective in separating the effects of housing price under different household structures.

Finally, even though this paper focuses on evidence from urban China, its basic mechanism can be applied to other countries as well. During Japan's housing bubble period in 1980s, there was a trend of rising household savings rate accompanied by rising coresidence rate. Adults living with older parents are quite common in Asian countries, such as Japan, Taiwan and South Korea. Many argued that the high coresidence rates in these countries are deeply rooted in the cultural background that filial piety or parental respect holds great value, however, in modern economy, the economic motivation can never be separated from social and cultural environment. This paper gives another perspective that integrates household structure and economic incentive.

Bibliography

Aghion, P., U. Akcigit, A. Bergeaud, R. Blundell, and D. Hemous (2017). Innovation and top income inequality. *NBER*. Working Paper 21247.

Akcigit, U., J. Grigsby, and T. Nicholas (2017). The rise of american ingenuity: Innovation and inventors of the golden age. *NBER*. Working Paper No. 23047.

Allen, T. and C. Arkolakis (2014). Trade and the topography of the spatial economy. *Quarterly Journal of Economics 129*, 1085–1140.

Arthur, W. (1994). *Increasing Returns and Path Dependence in the Economy*. University of Michigan Press.

Bacolod, M., B. S. Blum, and W. C. Strange (2009). Skills in the city. *Journal of Urban Economics 65*, 136–153.

Baldwin, R. and T. Okubo (2006). Heterogeneous firms, agglomeration and economic geography: Spatial selection and sorting. *Journal of Economic Geography 6*, 323–346.

Banerjee, A., X. Meng, and N. Qian (2010). The life cycle model and household savings: Micro evidence from urban china. Working paper.

Barth, J. R., M. Lea, and T. Li (2012). China's housing market: Is a bubble about to burst? *Milken Institute*.

Baum-Snow, N. and R. Pavan (2013). Inequality and city size. *Review of Economics and Statistics 95*, 1535–1548.

Behrens, K., G. Duranton, and F. Robert-Nicoud (2014). Productive cities: Sorting, selection, and agglomeration. *Journal of Political Economy 122*.

Behrens, K., G. Mion, Y. Murata, and J. Suedekum (2014). Trade, wages, and productivity. *International Economic Review 55*, 1305–1348.

Behrens, K. and F. Robert-Nicoud (2014). Survival of the fittest in cities: Urbanisation and inequality. *The Economic Journal 124*, 1371–1400.

Bleakley, H. and J. Lin (2012). Portage and path dependence. *Quarterly Journal of Economics 127*, 587–644.

Buzard, K. and G. Carlino (2013). *The Geography of Research and Development Activity in the U.S.*

Carlino, G. and W. R. Kerr (2014). Agglomeration and innovation. *NBER*. Working Paper No. 20367.

Carroll, C. D. (1997). Buffer-stock saving and the life cycle/permanent income hypothesis. *Quarterly Journal of Economics 112*(1), 1–56.

Chamon, M. and E. Prasad (2010). Why are saving rates of urban households in china rising. *American Economic Journal: Macroeconomics 2*(1), 93–130.

Chatterji, A., E. Glaeser, and W. Kerr (2014). Clusters of entrepreneurship and innovation. *NBER*. Working Paper No. w19013.

Chen, H., P. Gompers, A. Kovner, and J. Lerner (2010). Buy local? the geography of venture capital. *Journal of Urban Economics 67*, 90–102.

Chen, K., A. Imrohoroglu, and S. Imrohoroglu (2006). The japanese saving rate. *The American Economic Review 96*, 1850–1858.

Choukhmane, T., N. Coeurdacier, and K. Jin (2013). The one-child policy and household savings. *Mimeo, Science Po.*.

Combes, P.-P., G. Duranton, and L. Gobillon (2008). Spatial wage disparities: Sorting matters! *Journal of Urban Economics 63*, 723–742.

Combes, P.-P., G. Duranton, and L. Gobillon (2016). The costs of agglomeration: House and land prices in french cities. Working paper.

Curtis, C. C., S. Lugauer, and N. C. Mark (2015). Demographic patterns and household saving in china. *American Economic Journal: Macroeconomics 7*(2), 58–94.

Davis, D. R. and J. I. Dingel (2016). A spatial knowledge economy. *NBER*. Working Paper.

Davis, M. A. and S. V. Nieuwerburgh (2014). Housing, finance and the macroeconomy. *NBER*. Working Paper No. 20287.

Davis, M. A. and F. Ortalo-Magne (2011). Household expenditures, wages, rents. *Review of Economic Dy namics 14*, 248–261.

Deaton, A. and J. Muellbauer (1980). *Economics and Consumer Behavior*. Cambridge University Press.

Deaton, A. and C. H. Paxson (2000). Growth and saving among individuals and households. *Review of Economics and Statistics 82*(2), 212–225.

Desmet, K. and J. Henderson (2015). *The Geography of Development within Countries*, Volume 5, pp. 1457–1517.

Desmet, K. and E. Rossi-Hansberg (2013). Urban accounting and welfare. *American Economic Review 103*, 2296–2327.

Desmet, K. and E. Rossi-Hansberg (2014). Spatial development. *American Economic Review 104*, 1211–1243.

Diamond, R. (2016). The determinants and welfare implications of us workers' diverging location choices by skill: 1980-2000. *American Economic Review 106*, 479–524.

Duranton, G. (2007). Urban evolutions: the fast, the slow, and the still. *American Economic Review 97*, 197–221.

Duranton, G., J. V. Henderson, and W. C. Strange (2015). *Handbook of Regional & Urban Economics*, Volume 5.

Duranton, G. and D. Puga (2004). *Micro-foundations of Urban Agglomeration Economies*, Volume 4, pp. 2063–2117.

Eeckhout, J., R. Pinheiro, and K. Schmidheiny (2014). Spatial sorting. *Journal of Political Economy 122*, 554–620.

Esteve, A. and C. Liu (2013). Families in asia: A cross-national comparison of intergenerational coresidence.

Fajgelbaum, P. D. and A. K. Khandelwal (2016). Measuring the unequal gains from trade. *Quarterly Journal of Economics 131*, 1113–1180.

Fang, H., Q. Gu, W. Xiong, and L.-A. Zhou (2016). Demystifying the chinese housing boom. *NBER*. Working Paper.

Feldman, M. and D. Audretsch (1999). Innovation in cities: Science-based diver-

sity, specialization, and localized competition. *European Economic Review 43*, 409–429.

Gabaix, X., J.-M. Lasry, P.-L. Lions, and B. Moll (2016). The dynamics of inequality. *Econometrica 84*(6), 2071–2111.

Gale, W. G. and J. K. Scholz (1994). Intergenerational transfers and the accumulation of wealth. *Journal of Economic Perspectives 8*(4), 145–160.

Ganong, P. and D. Shoag (2016). Why has regional income convergence in the u.s. declined? *Journal of Urban Economics.* Working Paper.

Gaspar, J. and E. Glaeser (1996). Information technology and the future of cities. *NBER.* Working Paper.

Gennaioli, N., R. La Porta, F. Lopez-de Silanes, and A. Shleifer (2013). Human capital and regional development. *Quarterly Journal of Economics 128*, 105–164.

Giannone, E. (2017). Skilled-biased technical change and regional convergence. Working Paper.

Glaeser, E. (2011). *The Triumph of the City.* Penguin Books.

Glaeser, E. and J. D. Gottlieb (2009). The wealth of cities: Agglomeration economies and spatial equilibrium in the united states. *Journal of Economic Literature 47*, 983–1028.

Glaeser, E., M. Resseger, and K. Tobio (2009). Inequality in cities. *Journal of Regional Science 49*, 617–646.

Glaeser, E. and B. Ward (2009). The causes and consequences of land use regulation: Evidence from greater boston. *Journal of Urban Economics 65*, 265–278.

Goldin, C. and L. F. Katz (2007). Long-run changes in the wage structure: Narrowing, widening, polarizing. *Brookings Papers on Economic Activity* (2), 135–65.

Grossman, G. (2013). Heterogeneous workers and international trade. *Review of World Economics 149*, 211–245.

Gyourko, J., C. Mayer, and T. Sinai (2013). Superstar cities. *American Economic Journal: Economic Policy 5*, 167–99.

Henderson, J. V. (1974). The sizes and types of cities. *American Economic Review 64*, 640–56.

Henderson, V., T. Squires, A. Storeygard, and D. Weil (2017). The global distribution of economic activity: Nature, history, and the role of trade. *Quarterly Journal of Economics*. ISSN 0033-5533.

Holmes, T. J. and J. A. Schmitz (2010). Competition and productivity: A review of evidence. *Annual Reviews of Economics 2*(1), 619–642.

Hsieh, C.-T. and E. Moretti (2017). Housing constraints and spatial misallocation. *NBER*. Working Paper No. 21154.

Jones, C. I. and J. Kim (2017). A schumpeterian model of top income inequality. *Journal of Political Economy*. Forthcoming.

Kaplan, G. (2012). Moving back home: Insurance against labor market risk. *Journal of Political Economy 120*(3).

Kerr, W., R. Nanda, and M. Rhodes-Kropf (2014). Entrepreneurship as experimentation. *Journal of Economic Perspectives 28*, 25–48.

la Roca, J. D. and D. Puga (2017). Learning by working in big cities. *Review of Economic Studies 84*, 106–142.

Lee, J. and T. Nicholas (2013). *The Origins and Development of Silicon Valley.* Harvard Business School Case.

Li, B. and H. B. Shin (2013). Intergenerational housing support between retired old parents and their children in urban china. *Urban Studies 50*(16), 3325–3242.

Lucas, R. E. (2001). Externalities and cities. *Review of Economic Dynamics 4*, 245–274.

Lucas, R. E. and B. Moll (2014). Knowledge growth and the allocation of time. *Journal of Political Economy 122*, 1–51.

Manacorda, M. and E. Moretti (2006). Why do most italian youths live with their parents? intergenerational transfers and household structure. *Journal of the European Economic Association 4*(4), 800–829.

Melitz, M. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica 71*, 1695–1725.

Modigliani, F. and S. L. Cao (2004). The chinese saving puzzle and the life-cycle hypothesis. *Journal of Economic Literature 42*(1), 145–170.

Moretti, E. (2004). *Human Capital Externalities in Cities*, Chapter 51.

Moretti, E. (2012). *The New Geography of Jobs.* Houghton Mifflin Harcourt.

Perla, J. and C. Tonetti (2014). Equilibrium imitation and growth. *Journal of Political Economy, 122* (1), 52–76.

Perla, J., C. Tonetti, and M. E. Waugh (2016). Equilibrium technology diffusion, trade, and growth. *NBER*. Working paper.

Petersen, M. and R. Rajan (1994). The benefits of lending relationships: Evidence from small business data. *J. Finance 49*, 3–37.

Redding, S. J. (2016a). Goods trade, factor mobility and welfare. *Journal of International Economics 101*, 148–167.

Redding, S. J. and E. Rossi-Hansberg (2016b). Quantitative spatial economics. *NBER*. Working Paper No. 22655.

Rosenzweig, M. and J. Zhang (2014). Co-residence, life-cycle savings and intergenerational support in urban china. *NBER*. Working Paper.

Ruggles, S. and M. Heggeness (2008). Intergenerational coresidence in devel- oping countries. *Population and Development Review 34* (2), 253–281.

Simonovska, I. and M. Waugh (2014). The elasticity of trade: Estimates and evidence. *Journal of International Economics 92*, 34–50.

Song, Z. and D. T. Yang (2012). Life cycle earnings and saving in a fast-growing economy. Working Paper.

Tonetti, C. and J. Perla (2014). Equilibrium imitation and growth. *Journal of Applied Economics 18*, 122–148.

Wang, S.-Y. (2011). State misallocation and housing prices: Theory and evidence from china. *American Economic Review 101* (5), 2081–2107.

Waugh, M. E. (2010). International trade and income differences. *American Economic Review 100*, 2093–2124.

Wei, S.-J. and X. Zhang (2011). The competitive saving motive: Evidence from rising sex ratios and savings rates in china. *Journal of Political Economy 119*(3), 511–564.

Wu, J., J. Gyourko, and Y. Deng (2010). Evaluating conditions in major chinese housing markets. *NBER*. Working Paper.

APPENDIX A

APPENDIX A1

Proof of proposition 1: I begin by solving the final goods sector's problem. Since the final good is perfectly competitive and its price is normalized to 1, given the price p_{ji} of intermediate good i in city j, the following problem solves for the demand of intermediate good i :

$$\max_{y_{ji}} \left(\int_{i \in \Omega_j} y_{ji}^{\theta} di \right)^{1/\theta} - \int_{i \in \Omega_j} p_{ji} y_{ji} di;$$

The demand equations of variety i derived from first order conditions are

$$p_{ji} = \left(\frac{Y_j}{y_{ji}} \right)^{1-\theta}. \tag{2.7.1}$$

According to the intermediate goods production function (1.3.3), the marginal cost of intermediate goods production is $\frac{w_j}{q_{ji}^{\gamma}}$, then profit maximization yields the optimal pricing rule

$$p_{ji} = \frac{1}{\theta} \frac{w_j}{q_{ji}^{\gamma}}, \tag{2.7.2}$$

where $\frac{1}{\theta}$ equals the mark-up factor. Together with (A.1) an (A.2), the intermediate good output y_{ji} is

$$y_{ji} = \left(\frac{w_j}{\theta q_{ji}^{\gamma}} \right)^{\frac{1}{\theta-1}} Y_j. \tag{2.7.3}$$

Replacing $y_{ji} = \left(\frac{w_j}{\theta q_{ji}^{\gamma}} \right)^{1/(\theta-1)} Y_j$ into final good production function (1.3.3), the wage rate of city j is:

$$w_j = \theta Q_j^\gamma \qquad (2.7.4)$$

This expression implies that the equilibrium wage increases with city-specific productivity, thus the city with more skilled entrepreneurs and more intensive innovation activities has a higher wage. Rewriting the intermediate good demand function (A.3) with wage equation (A.4), the output of intermediate good is

$$y_{ji} = \left(\frac{q_{ji}}{Q_j}\right)^{1/\theta} Y_j. \qquad (2.7.5)$$

Combine final good production function and (A.5), together with the labor market clearing condition $L_j = \int_{i \in \Omega_j} l_{ji} di$, the aggregate output in city j is

$$Y_j = Q_j^\gamma L_j. \qquad (2.7.6)$$

Now it is time to solve for entrepreneur's profit function. Taking advantage of (A.2), (A.3) , (A.4) and (1.4.6), the expected profit from selling variety for entrepreneur i in city j is:

$$E\left(\pi_{ji} \mid z\right) = \frac{1-\theta}{\theta} \frac{w_j}{q_{ji}^\gamma} \left(\frac{w_j}{\theta q_{ji}^\gamma}\right)^{\frac{1}{\theta-1}} Y_j = (1-\theta) Y_j \frac{z}{Q_j} \left(\frac{\phi_{ji} + \alpha - 1}{\alpha - 1}\right). \qquad (2.7.7)$$

Hence proposition 1 is proved.

APPENDIX A2

Proof of proposition 2: Suppose when an entrepreneur's productivity level is equal to $\bar{z}$, his indirect utility from living in city 1 and city 2 are the same,

meaning

$$\Psi_1 \bar{z} - a_1 = \Psi_2 \bar{z} - a_2,$$

For any constant number $\lambda > 0$, the following equation holds:

$$\lambda \bar{z} \left(\Psi_1 - \Psi_2 \right) = \lambda \left(a_1 - a_2 \right).$$

If $\lambda > 1$, then there is

$$\lambda \bar{z} \left(\Psi_1 - \Psi_2 \right) > \left(a_1 - a_2 \right)$$

$$\Rightarrow \Psi_1 \left(\lambda \bar{z} \right) - a_1 > \Psi_2 \left(\lambda \bar{z} \right) - a_2.$$

which means when an entrepreneur's productivity is $\lambda \bar{z}$, for any $\lambda > 1$, his net profit in city 1 is higher than that of city 2, thus he will choose to locate in the big city. Similarly, if $0 < \lambda < 1$, then for an entrepreneur with productivity of $\lambda \bar{z}$, his net profit in city 1 is smaller than that of city 2, thus he will locate at the small city.

APPENDIX A3

Proof of proposition 4: According to section 1.2.7, the skill sorting threshold is $\bar{z} = \frac{a_1 - a_2}{\Psi_1(Q_1) - \Psi_2(Q_2)}$, with

$$\Psi_j \left(Q_j \right) = \frac{1}{\eta_j} \left(\frac{\theta}{u^* + a_j} \right)^{\frac{\mu}{1-\mu}} \left[\frac{\gamma \alpha}{(\alpha - 1)} - 2 \left(\frac{\gamma \chi_j}{(\alpha - 1)} \right)^{1/2} \right] Q_j^{\frac{\gamma \mu + \mu - 1}{1-\mu}} = \Lambda_j Q_j^{\frac{\gamma \mu + \mu - 1}{1-\mu}}.$$

In addition, equation (1.4.16) suggests that city $j's$ productivity is an increasing function of the skill sorting threshold $\bar{z}$, such that $\frac{\partial Q_i(\bar{z})}{\partial \bar{z}} > 0$. Rewrite the skill sorting threshold expression:

$$\bar{z} = \frac{a_1 - a_2}{\Lambda_1 Q_1^{\frac{\gamma\mu+\mu-1}{1-\mu}} - \Lambda_2 Q_2^{\frac{\gamma\mu+\mu-1}{1-\mu}}}$$

$$= \frac{1}{(Q_2(\bar{z}))^{\frac{\gamma\mu+\mu-1}{1-\mu}}} \frac{a_1 - a_2}{\Lambda_1\left(\frac{Q_1}{Q_2}(\bar{z})\right)^{\frac{\gamma\mu+\mu-1}{1-\mu}} - \Lambda_2}. \qquad (2.7.8)$$

We know from (1.4.16) that the relative productivity $\frac{Q_1}{Q_2}$ is an increasing function of the skill sorting threshold $\bar{z}$: $\frac{\partial\left(\frac{Q_1}{Q_2}(z)\right)}{\partial \bar{z}} > 0$. Then expression (A.8) indicates that threshold $\bar{z}$ increases in aggregate distribution's minimum productivity threshold z_{min}, such that $\frac{\partial \bar{z}}{\partial z_{min}} > 0$. Hence the relative productivity is also increasing in minimum productivity threshold z_{min}:

$$\frac{\partial\left(\frac{Q_1}{Q_2}\right)}{\partial z_{min}} > 0.$$

This positive relationship indicates that TFP growth reinforces the skill sorting process, making the big city relatively more productive than the small city. Nest, I analyze the effects of locational fundamentals on skill sorting threshold. Since $\Lambda_1 = \frac{1}{\eta_1}\left(\frac{\theta}{u^*+a_1}\right)^{\frac{\mu}{1-\mu}}\left[\frac{\gamma\alpha}{(\alpha-1)} - 2\left(\frac{\gamma\chi_1}{(\alpha-1)}\right)^{1/2}\right]$ is decreasing in χ_1, η_1 and a_1, and it is independent in $\bar{z}$, then equation (A.8) can generate the following results:

$$\frac{\partial \bar{z}}{\partial \chi_1} > 0, \quad \frac{\partial \bar{z}}{\partial \eta_1} > 0 \; and \; \frac{\partial \bar{z}}{\partial a_1} > 0.$$

A higher threshold $\bar{z}$ indicates stronger skill sorting effects, and a higher relative productivity $\frac{Q_1}{Q_2}$ accordingly. Meanwhile the city-specific wage rate is an increasing function of city productivity, see (1.4.3). Therefore, the relative wage rate between

the two cities $\frac{w_1}{w_2}$ also rises with the sorting threshold. In summation:

$$\frac{Q_1}{Q_2} \uparrow \ and \ \frac{w_1}{w_2} \uparrow \ if \ \chi_1 \uparrow \ or \ \eta_1 \uparrow \ or \ a_1 \uparrow$$

Now, let's examine the change in relative population density $\frac{L_1/\bar{H}_1}{L_2/\bar{H}_2}$:

$$\frac{L_1/\bar{H}_1}{L_2/\bar{H}_2} = \left(\frac{w_1}{w_2}\right)^{\frac{\mu}{1-\mu}} \frac{(u^* + a_2)^{\frac{1}{1-\mu}}}{(u^* + a_1)^{\frac{1}{1-\mu}}}. \tag{2.7.9}$$

We know an increase in χ_1 or η_1 lead to an increase in $\frac{w_1}{w_2}$, then the relative population density $\frac{L_1/\bar{H}_1}{L_2/\bar{H}_2}$ also increases with an increase in χ_1 or η_1. But the effect of a_1 on $\frac{L_1/\bar{H}_1}{L_2/\bar{H}_2}$ is ambiguous.

APPENDIX A4

Proof of proposition 6: According to section 1.3, the skill sorting threshold is $\bar{z}^* = \frac{a_1 - a_2}{\tilde{\Psi}_1(Q_1) - \tilde{\Psi}_2(Q_1)}$, with

$$\tilde{\Psi}_j(Q_j) = \frac{\gamma}{\eta_j} \left(\frac{\theta}{u^* + a_j}\right)^{\frac{\mu}{1-\mu}} \left[\frac{\alpha - 1 + \beta_j}{(\alpha - 1)} - 2\left(\frac{\chi\beta_j}{\gamma(\alpha - 1)}\right)^{\frac{1}{2}}\right] Q_j^{\frac{\gamma\mu + \mu - 1}{1-\mu}}.$$

where $\beta_j = \frac{\bar{z}_j}{\bar{z}_j + c}$ stands for the local learning opportunity. For simplicity, let

$$\tilde{\Psi}_j(Q_j) = \tilde{\Lambda}_j Q_j^{\frac{\gamma\mu + \mu - 1}{1-\mu}}.$$

Then simple algebra indicates that $\frac{\partial \tilde{\Lambda}_j}{\partial \beta_j} > 0$. In addition, equation (1.5.5) suggests that city $j's$ average productivity is an increasing function of the skill sorting threshold $\bar{z}^*$, such that $\frac{\partial Q_i(\bar{z}^*)}{\partial \bar{z}^*} > 0$. Furthermore, the relative productivity $\frac{Q_1}{Q_2}$ is

also an increasing function of the skill sorting threshold $\bar{z}^*$:

$$\frac{\partial \left(\frac{Q_1}{Q_2} (\bar{z}^*) \right)}{\partial \bar{z}^*} > 0.$$

Rewrite the skill sorting threshold expression:

$$\bar{z}^* = \frac{a_1 - a_2}{\tilde{\Lambda}_1 Q_1^{\frac{\gamma\mu+\mu-1}{1-\mu}} - \tilde{\Lambda}_2 Q_2^{\frac{\gamma\mu+\mu-1}{1-\mu}}}$$

$$= \frac{1}{Q_2^{\frac{\gamma\mu+\mu-1}{1-\mu}}} \frac{a_1 - a_2}{\tilde{\Lambda}_1 \frac{Q_1}{Q_2}^{\frac{\gamma\mu+\mu-1}{1-\mu}} - \tilde{\Lambda}_2}. \tag{2.7.10}$$

If information technology and cities are complement, then a drop in c will lead to a greater improvement in big city's learning opportunity, such that $\triangle\beta_1 > \triangle\beta_2$. Therefore, there is $\triangle\tilde{\Lambda}_1 > \triangle\tilde{\Lambda}_2$ in this case, which means the equilibrium skill sorting threshold decreases as c decreases:

$$\frac{\partial \bar{z}^*}{\partial c} > 0.$$

If information technology and cities are substitute, then a drop in c will lead to a greater improvement in small city's learning opportunity, such that $\triangle\beta_1 < \triangle\beta_2$. Therefore, there is $\triangle\tilde{\Lambda}_1 < \triangle\tilde{\Lambda}_2$ in this case, which means the equilibrium skill sorting threshold increases as c decreases:

$$\frac{\partial \bar{z}^*}{\partial c} < 0.$$

The above differentiated results suggest that information technology can cause substantial changes to city's skill compositions. Because the way people interact with each other as well as its relationship with geography can be quite important in terms of economic development and economic activity distribution across space.

Appendix A5

Proof of proposition 7: The goods market clearance condition imply the local aggregate output Y_j is the sum of labor income $w_j L_j$, entrepreneurial profit from domestic market $(1-\theta)\left(\frac{\theta}{w_j}\right)^{1/\gamma} Q_j Y_j$ and export profit $(1-\theta)\left(\frac{1}{\tau_{jn}}\right)^{1/\gamma}\left(\frac{\theta}{w_j}\right)^{1/\gamma} Q_j Y_n$, hence we have

$$(1-\theta)\left(\frac{\theta}{w_j}\right)^{1/\gamma} Q_j \left[Y_j + \left(\frac{1}{\tau_{jn}}\right)^{1/\gamma} Y_n\right] + w_j L_j = Y_j; \tag{2.7.11}$$

With factor market clearance condition (1.6.4), we can establish the relationship between city-specific wage rate and average city productivity:

$$\left(\frac{w_j}{\theta}\right)^{\frac{1}{\gamma}} = Q_j \left[1 + \left(\frac{1}{\tau_{jn}}\right)^{\frac{1}{\gamma}} \frac{(\tau_{jn})^{\frac{1}{\gamma}} - 1}{(\tau_{nj})^{\frac{1}{\gamma}} - 1}\right]. \tag{2.7.12}$$

Then the trade balance condition (1.6.6), goods market clearance condition (A.9) and the wage equation (A.10) lead to the following conclusions:

$$w_j L_j = \theta Y_j \tag{2.7.13}$$

$$\frac{Y_1}{Y_2} = \frac{w_1 L_1}{w_2 L_2} = \frac{(\tau_{21})^{\frac{1}{\gamma}} - 1}{(\tau_{12})^{\frac{1}{\gamma}} - 1}. \tag{2.7.14}$$

Assuming asymmetric trade cost $\tau_{21} > \tau_{12}$, if there is a proportional drop in trade cost (namely $\hat{\tau}_{12} = \rho\tau_{12}$, $\hat{\tau}_{21} = \rho\tau_{21}$ and $0 < \rho < 1$), then the relative economy size becomes larger in response to the trade cost decline

$$\frac{\hat{Y}_2}{\hat{Y}_1} = \frac{(\hat{\tau}_{21})^{\frac{1}{\gamma}} - 1}{(\hat{\tau}_{12})^{\frac{1}{\gamma}} - 1} > \frac{(\tau_{21})^{\frac{1}{\gamma}} - 1}{(\tau_{12})^{\frac{1}{\gamma}} - 1}. \tag{2.7.15}$$

Due to the positive relationship between wage and productivity described in (A.12), and the positive link between city size and wage rate presented in (1.4.12), the relative productivity can be expressed as:

$$\frac{Q_1}{Q_2} = \frac{(\tau_{12})^{\frac{1}{\gamma}}}{(\tau_{21})^{\frac{1}{\gamma}}} \left[\frac{\eta_1 (u_X^* + a_1)}{\eta_2 (u_X^* + a_2)} \right]^{\frac{1}{\gamma}} \left[\frac{(\tau_{21})^{\frac{1}{\gamma}} - 1}{(\tau_{12})^{\frac{1}{\gamma}} - 1} \right]^{\frac{1-\mu}{\gamma}+1}. \tag{2.7.16}$$

where u_X^* is the equilibrium common utility of workers under the trade model, such that $\frac{w_1}{p_{1j}^{1-\mu}} - a_1 = \frac{w_2}{p_{h2}^{1-\mu}} - a_2 = u_X^*$. Taking account of (A.15) and (A.16), there is

$$\frac{\hat{Q}_1}{\hat{Q}_2} = \frac{(\hat{\tau}_{12})^{\frac{1}{\gamma}}}{(\hat{\tau}_{21})^{\frac{1}{\gamma}}} \left[\frac{\eta_1 (u_X^* + a_1)}{\eta_2 (u_X^* + a_2)} \right]^{\frac{1}{\gamma}} \left[\frac{(\hat{\tau}_{21})^{\frac{1}{\gamma}} - 1}{(\hat{\tau}_{12})^{\frac{1}{\gamma}} - 1} \right]^{\frac{1-\mu}{\gamma}+1} > \frac{Q_1}{Q_2}. \tag{2.7.17}$$

Therefore, the relative productivity between the large and small city increases as the trade cost decreases. Since the relative productivity is increasing in the skill sorting threshold $\bar{z}^X$, the negative relationship between skill sorting threshold $\bar{z}^X$ and trade cost is thereby proved. In addition, the equilibrium relative wage rate $\frac{w_1}{w_2}$ also increases with a decrease in trade costs.

$$\frac{w_1}{w_2} = \frac{\eta_1 (u_X^* + a_1)}{\eta_2 (u_X^* + a_2)} \left[\frac{(\tau_{21})^{\frac{1}{\gamma}} - 1}{(\tau_{12})^{\frac{1}{\gamma}} - 1} \right]^{1-\mu}. \tag{2.7.18}$$

Equation (A.18) indicates that if trade costs decline, the new equilibrium relative wage $\frac{\hat{w}_1}{\hat{w}_2}$ is

$$\frac{\hat{w}_1}{\hat{w}_2} = \frac{\eta_1 (u_X^* + a_1)}{\eta_2 (u_X^* + a_2)} \left[\frac{(\hat{\tau}_{21})^{\frac{1}{\gamma}} - 1}{(\hat{\tau}_{12})^{\frac{1}{\gamma}} - 1} \right]^{1-\mu} > \frac{w_1}{w_2}. \tag{2.7.19}$$

The above proof means that when trade costs decline by the same percentage in the two cities, the sorting threshold increases accordingly, so are the relative productivity and relative wage.