Essays on Market Structure

Feng Ruan

Doctor of Philosophy in Economics

The University of Edinburgh

2011

Acknowledgements

I have been incredibly lucky to have worked under the guidance and supports of József Sákovics and Ahmed Anwar. I am profoundly indebted to my primary supervisor József Sákovics for all he has taught me. He is extremely knowledgeable about the subject and exceptionally generous with his time giving me advice. His vision of economics has greatly influenced the way I think of the subject. Ahmed Anwar sets a perfect example as a scholar. I appreciate his inspiring comments on my study and encouragement.

I am grateful to the School of Economics for providing me with an extremely stimulating environment. I also own my gratitude to Ed Hopkins for giving me the opportunity of doing a PhD in the University of Edinburgh. I would like to thank Kohei Kawamura, Nick Vikander, and Jakub Steiner for the insightful discussions during my studies.

I would also like to thank China Scholarship Council and Shanghai University of Finance and Economics for the financial support during the past three years.

Last but not least, I would like to thank my parents, Weiguo Ruan and Xiujun Shi, for their love and constant supports on my studies. Without their sacrifice and supports, I cannot even take the first step in my overseas studies. Also special thanks to Xiaoxiao for always being a perfect audience and not thinking that I was wasting my time.

Declaration

I declare that this thesis is the result of my own research except as cited in the references.

The thesis has not been accepted for any degree and is not concurrently submitted in candidature of any other degree.

Feng Ruan, November, 2011

Contents

List of Figures

Chapter 1

Introduction

Some of the most important work in the development of economic theory is associated with the study of market structure. In essence, most markets are two-sided. For example, product markets connect tens of thousands of product brands to tens of millions of consumers; marriage markets couple the single men and women who would otherwise suffer from a lonely heart; and labour markets link the job candidates to their preferred employers and positions. Apart from the two-sidedness, we have explored another important common aspect of these market structures, i.e. interconnection/competition of the segments within one side of the market. Under this common thread, the three essays in this thesis are freshly formulated in a loosely related manner, covering topics in three different areas.

Chapter 2 is motivated by strategic transitions of many marketplaces (e.g. Amazon.com). From the perspective of a platform owner, when it owns part of the business on one side of the market, there is no straightforward answer as to whether having the rest of business owned by others is advantageous or not. The argument is that, on the one hand, the platform welcomes more third-party business as it boosts revenue in terms of membership fees; on the other hand the business owned by the platform dislikes the incoming competitors whose participation drives down profit margins. We propose a novel framework in this chapter to explore the trade-off between the two. Here, the intermediary can decide to be either a "merchant" or a "two-sided platform", or a hybrid one in between. Our analysis shows that in hybrid mode the platform extracts all the surplus from the producers of the merchandised brands, and the merchandised brands always charge a price premium compared to the directly retailed ones. We also show that as the platform absorbs an existing directly retailed brand into the self-brand portfolio, the equilibrium prices of both brand types are increased. We find that only the directly retailed brands dominate the market when the platform's capacity is relatively small; and both brand types coexist in the marketplace when the capacity is relatively large. Furthermore, we find a backward bending proportion plus a vertical proportion of the "contract curve" in comparative statics. That is, the self-brand portfolio always expands while the third-party-brand portfolio shrinks until it reaches a certain level, when the platform increases its capacity. It helps us to gain some ideas on the dynamics of brand portfolio management for the platform. Lastly, taking into account of indirect network effect which is the common feature in the two-sided market, it is shown that the platform is better off when consumers have positive expected surplus.

Chapter 3 is much motivated by the Chinese experience. China has witnessed the largest rural to urban labour flow (among which the majority are male) in the world's history over the last three decades. We propose an idea that the grand migration can also be attributed to the unbalanced sex ratio between rural and urban areas. This chapter develops a two-sided matching model of two linked marriage markets with homogeneous agents, non-transferable utility and search friction. We extend the one-market model of the previous literature into a two-market one, allowing the agents to migrate between the markets at a fixed cost. The analysis focuses on the unmatched as well as the migrating population, which is induced by the different sex ratios in the two geographically isolated marriage markets. We find that imperfections in the matching technology leads to the enlarged gap of sex ratio of the unmatched population compared to that of the unbalanced inflows. We are interested in the question of how the migrating costs affect the migration between rural and urban areas, and under what conditions a subsidy covering migrating costs might benefit a party in the marriage markets. We characterise the equilibrium set in the parameter space of migrating costs, and find that a full subsidy of migrating costs does not necessarily benefit those who receive it but always benefits the opposite sex, if they are the short sides of both markets.

Chapter 4 explains the migration of labour force from a different angle. Here, the migration is of workers to jobs. Motivated by the distinction of public and private sector, we consider a spatial oligopsony model in which firms (two co-locating small firms with recruiting capacity constraints and a large firm without such limit) are competing for workers along a "strip" market. The capacity issue that is extensively discussed in the Chapter 2 again plays an important role in this model, though in a very different context. It is shown that the recruiting capacity affects the intra-group competition and hence the inter-group competition in wage-posting strategies. Additionally, we show that, as recruiting limits expand, the expected wages offered by the small firms increase while the wage offered by the big firm decreases, which helps to explain the recent trend of the wage disparity between public and private jobs. We also characterise the equilibrium wages and the size (direction) of the migration in the three-stage game (i.e. the workers decide whether to relocate in the first stage, then the big firm decides its wage offer, and lastly, the two co-locating firms simultaneous set wages), which helps us to understand better the inter-sector mobility in a changing environment of economy.

We investigate the issues of interconnection and competition in three different markets. It is always of interest for a researcher of economics to have some ideas on the same issue from different perspectives. Remember that whilst this is a collection of essays on economic theory, it is nonetheless compared to empirical observation. And it will surely serve as a starting point for the author to further the research on market structure.

Chapter 2

Brand Portfolio in Two-sided Market: Platform Merchandised Brands versus Directly Retailed Brands

Abstract

We examine the hybrid business model for an intermediary acting as both a "merchant" and a "two-sided platform". The dual role of the platform naturally divides the brands into two groups: platform merchandised brands and directly retailed brands. While the platform makes a profit by merchandising and rent-collecting from the affiliated direct retailers, we explicitly explore the trade-offs between the two. Our analysis shows that in a hybrid mode the platform extracts all the surpluses from the producers of the merchandised brands, and the merchandised brands are always priced higher than the directly retailed ones. We also show that, as the platform absorbs an existing directly retailed brand into the self-brand portfolio, the equilibrium prices of both brand types are increased. We find that when the platform's capacity is relatively small, only the directly retailed brands dominate the market; and when the capacity is relatively large both brand types coexist in the marketplace.

Furthermore, we find a backward bending proportion plus a vertical proportion of the "contract curve" in comparative statics analysis. That is, when the platform increases its capacity, the self-brand portfolio always expands, whereas the third-party-brand portfolio shrinks until it reaches a certain level. When there is indirect network effect in consumer demand and positive expected surplus for consumers, the platform is better off; whereas the direct retailers are worse off.

2.1 Introduction

In our economy, consumers are becoming increasingly dependent on one-stop shopping for its unbeatable features of cost-efficiency and convenience. For traditional shopping, many people will still visit the same nearby supermarket chain (e.g. Tesco, Sainsbury's) or the same shopping mall for daily/weekly purchases. For those who feel more comfortable shopping in pajamas at the desk, they regularly go to the online stores (e.g. Tesco Direct, Amazon.com, eBay, Argos) in search of the goods they want to consume. In all these cases, the consumers must join a certain market platform in order to gain the access to the final goods.

In this sense, it is the market platform that connects the tens of millions of consumers to the tens of thousands of brands. Taking a closer look, the market intermediation can be classified into two forms (Hagiu, 2007):

(1) "Intermediary as a middleman (or merchant)", which buys goods from sellers and resells to buyers, e.g. supermarket chains and other traditional stores. They make a profit by exploiting the price difference between wholesaling and retailing.

(2) "Intermediary as a two-sided platform", which allows third-party retailers to sell directly to the buyers visiting the same platform. Shopping malls and eBay are the most prominent examples, and they profit by collecting membership fees (or rents) from these affiliated direct retailers.

However, a number of market organisations have undertaken the transition from a pure "merchant" or "two-sided platform" to a hybrid one. For example, Amazon.com initially started as an online bookstore, but it has expanded the range of its merchandise at a surprising rate over the past few years. More interestingly, meanwhile, it started to allow the independent direct retailers to do business on the same platform (which is called Amazon Marketplace) at a fee. Another digital giant, Apple, on the one hand has expanded its iTunes products from the ¢99 song to the $1.99 TV episode ($2.99 for an HD version), and similar with many other applications; and on the other hand accommodates a huge variety of third-party developers who sell their own applications in the same online App Store. Moreover, a large proportion of the traditional supermarket sector now both accommodate the directly retailed brands on the shelves at some rental fee, whilst also absorbing more product lines into their self-brand portfolio[1] (e.g. "Finest" range for Tesco and "Taste the Difference" range for Sainsbury's). In this chapter, we will show the rationale behind this trend from the perspective of the platform owners.

In the case of the hybrid business model, the platform's profits come from two main sources: (1) the profits from wholesaling and retailing the merchandised brands (also referred to as the self-brand portfolio); (2) the rents collected from the affiliated direct retailers on the platform.

[1] This is also referred to as Category or Brand Management in marketing literature.

Unlike the pure "merchant" or the pure "two-sided platform" business model, the dual role of the hybrid platform naturally divides the brands into two groups (i.e. the self-brand portfolio and the third-party-brand portfolio), and there is potential competing conflict between the two. Thus, there is no straightforward answer as to whether the directly retailed brand is friend or foe to the platform.

In many cases, neither the pure "merchant" nor the pure "two-sided platform" business model is optimal for the platform if the intermediary choice in between is an option. The intuition is straightforward: As a pure "merchant", the platform might have the incentive to rent out existing or additional slots to the direct retailers for more profits. As a pure "two-sided platform", the intermediary might be lured by the high profit margin earned by the individual retailers, and engage itself as "merchant" for a select range of products. It is still far from clear what a platform's optimal strategy might look like, and it will soon lose its tractability if we focus only on the direct choice variables (i.e. the wholesaling prices and the admission fees.)

Using a simple model in this chapter, we claim that the platform effectively controls the brand portfolio composite by setting the wholesale price and the admission fees. Without losing sight of the whole picture, we take steps to address what the optimal brand portfolio is for the platform in this context.

In the first subsection, we discuss the case where the platform is restricted to host only a certain variety of brands. We formalise the trade-offs between the "intermediary-as-a-middleman" and the "intermediary-as-a-two-sided-platform" business model from the perspective of the platform owner. We show that the pure "two-sided platform" mode dominates the hybrid mode when the platform's hosting capacity is under a certain level; and we show that both brand types coexist on the platform when its hosting capacity is above a certain level.

We give a detailed discussion on the hybrid model, and show that the platform merchandised brands are always priced higher than the directly retailed ones. Additionally, the platform always bids for the merchandises at the marginal cost, and hence extracts all the surpluses from the producers of the merchandised brands. We also examine the platform's incentive for absorbing one more merchandised brand at a fixed level of brand variety. Our analysis shows that, as the platform includes an existing directly retailed brand into its own self-brand portfolio, the equilibrium prices of both brand types are increased. It is beneficial to the platform in a sense that its self-brand portfolio now earns a higher profit margin, and also it can extract more rent from the affiliated direct retailers who now similarly earn a higher profit margin. However, the hybrid platform will not endlessly absorb all the brands into its self-brand portfolio since the platform has to balance the negative impact on demand that is brought by the rising price level.

In the second subsection, we are interested in comparative statics with respect to the platform's accommodating capacity since it captures some intuition about the dynamics of the optimal brand portfolio as the scale of the platform expands. It is interesting to find that when the intermediary increases its capacity the self-brand portfolio always expands, whereas the third-party-brand portfolio shrinks until it reaches a certain level.

In the third subsection, we investigate how indirect network effect in consumer demand affects our previous results. It is shown that, if the expected surplus for the consumer is positive, the indirect network effect makes the platform owner better off but the direct retailers worse off.

Finally, we examine how the value of the outside option affects the optimal brand mix at every level of brand variety. We find that the pure "merchant" mode is preferred only when the outside option is not at all attractive.

2.1.1 Literature Review

Since the proposed framework in this chapter connects the pure "intermediary-as-a-middleman" and the pure "intermediary-as-a-two-sided-platform" model, there are essentially three separate strands of economics research contributing to this.

There is rich literature on the micro-structure of the market organisation with middlemen intermediating between buyers and sellers. Rubinstein and Wolinsky (1987) propose that the middleman who buys a good from one individual and sells to another at a higher price can be active in market when competing with the possible direct exchange in a random matching process. Spulber (1996) allows competition among middlemen when buyers and sellers search across them. Rust and Hall (2001) distinguish between the two competing types of intermediaries: "middlemen" (dealers/brokers) – whose price quotes are private information and can only be obtained by costly search process – and "market-makers" (specialists) – who post publicly observable bid and ask prices. We assume in this chapter that all the prices quoted are public information but do not distinguish the use of terminology of either middleman or market-makers; instead, we call them "merchants". Additionally, we assume that, on the one hand, there is no difference in search cost for the consumers once the products are brought on board by the middleman or by the direct retailers; and on the other hand, the search cost is punitively high for the consumer if he/she stays off board.

There is also burgeoning literature on the intermediation of two-sided platform, which emphasises the indirect network effects. In this type of market, the agents on each side of the market benefit from joining the same platform and interacting with the other side. Therefore the platform can make profit by charging from the agents on both sides of the market. Rochet and Tirole (2006) distinguish between membership charges and usage charges, and provide a model integrating the externalities exerted by the two and unifying these two seemingly different strands of literature. Rochet and Tirole (2003) and Armstrong (2006) explain the pricing structure of a monopoly platform and also examine the competition amongst the two-sided platforms. Virtually all the papers on the two-sided market focus on the pricing structure decided by the platform. In this chapter we consider a monopolist platform model but only focus on the subscription fee imposed on the affiliated direct retailers, rather than buyers[2]. Additionally, we look at the bidding price set by the platform for the merchandised brands, as we mainly discuss the hybrid mode.

Furthermore, there is one additional strand of literature that tries to uncover the trade-offs between the above mentioned two distinctive business models. The most closely related paper is "Merchant or Two-Sided Platform?" by Hagiu (2007). He compares two polar strategies for market intermediation which he names "merchant" mode and "two-sided platform" mode. While most of the literature on the two-sided market takes the indirect network effect as given, he argues that the belief on realisation of this effect is crucial to retailers' decisions of whether or not to affiliate. Consequently, the merchant mode is always preferred as the indirect network externalities will be fully internalised by the intermediary's buy-out contract. Hagiu also claims that the investment incentives and asymmetric information might lead to the opposite optimal strategy: being a pure "two-sided platform". Therefore, the strategic decision for the intermediary is not a 0-1 notion but a position along this continuum. In spite of the similar result of a hybrid optimal strategy, our model differs from theirs in that we endogenise the participating decision for the sellers by allowing them to choose between simply selling to the middleman or directly to consumers. This is a crucial point in this chapter. More importantly, without the additional assumptions in Hagiu's model, we argue that the existence of the optimal hybrid strategy for the intermediary being both a middleman and a two-sided platform should be mainly attributed to the presence of the outside option for the producers.

There is other relevant literature that bears similar intuition to this chapter. Nocke, et. al. (2007) develop a model comparing the impact of monopoly and dispersed ownership of the platform. The former coincides with the conventional setup of a monopoly two-sided platform, while in the latter case the platform ownership has been assigned to a certain group of sellers who have the vote to decide the admission fee and effectively influence the entry of other sellers. It is interesting to see that their model and ours are in fact two sides of the same coin. That is, their model is drawn from the viewpoint of the dispersed owners who are entitled to a fixed proportion of the market; while ours is from the perspective of a monopoly platform owner who is flexible in deciding any combination of merchandised brands and directly retailed brands in its portfolio. Additionally, this chapter is related to the literature on competition among a multi-product firm and many single-product firms in a differentiated market (e.g. Giraud-Heraud et. al. ,2003), despite the fact that in our model the "multi-product firm" (or the platform) also collects rental fees from other single-product firms and has to internalise the

[2] We assume zero access price to the consumers since it is not feasible to charge buyers upon entry for most of the retailing business.

relevant externalities.

The plan of the chapter is as follows. The basic model is presented in Section 2.2, and the results of the equilibrium analysis are shown in Section 2.3. There we discuss the existence of the optimal brand portfolio, and analyse the platform's incentive to absorb one more existing directly retailed brand at a fixed level of brand variety (Subsection 2.3.1). We show the results of the comparative statics with respect to the intermediary's accommodating capacity in Subsection 2.3.2; and we discuss how the indirect network effect in consumer demand and the value of the outside option affect our results in Subsection 2.3.3 and Subsection 2.3.4. Section 2.4 concludes.

2.2 The Model

It is a product market that is played by three types of agents: platform, sellers and buyers.

The products: The products are horizontally differentiated in an infinitely large menu of brands indexed by i, $i = 1, 2, ..., \infty$. For brand i, it incurs a distribution cost of $g(i)$ to make its goods available in the market. If we list the menu of brands in the ascending order of the distribution cost, $g(i)$ is an increasing function that is common knowledge to all, and is independent of demand. Additionally, there is a constant marginal cost c common to all brands.

The platform: In a two-sided market, the platform normally charges positive or negative[3] subscription fees to the agents on either side of the market. For simplicity, we ignore the admission fee imposed on the consumers and restrict ourselves to the choice of charging only membership fee r to sellers. In addition to the conventional setup in the two-sided market, we formulate that the platform can join the retailing channel by merchandising; that is, the platform purchases brand i from the seller at a universal wholesale price p^w and then sells to the buyers at price p_i^o; however, at the same time it has to take care of the relevant distribution cost $g(i)$ for the merchandised brand i.[4] Meanwhile, the platform collects the membership fee r from the affiliated direct retailers who do business on the platform and cover the distribution costs themselves.

Buyers: There is a continuum of buyers with unit measure. The buyers pay a visit to the platform for the end products which are horizontally differentiated in brands. Each buyer can choose to purchase only a single unit of one of the brands, or not. The utility for a typical buyer l purchasing brand i is $v_i^l = v - p_i + \varepsilon_i^l$. v is the systematic value that each buyer receives from consuming any of the brands; p_i is the price per product for brand i; ε_i^l denotes the idiosyncratic

[3]Negative tariffs can be subsidies.

[4]We assume that the platform is no more efficient than the individual direct retailers in dealing with the distribution costs. There are cases where the wholesaler is more efficient due to the economies of scale; these are not discussed in this paper.

taste that the buyer l has about brand i.[5] The buyer decides to purchase one product of brand i if and only if its net utility exceeds that of all other brands and exceeds zero, (which is the utility of non-purchasing). Thus, the mass of consumers who finally make purchases can be denoted as $D(\mathbf{p})$ where $\mathbf{p}$ is the vector of the market prices.

Sellers: Each brand is controlled by a unique seller. There are two ways (i.e. Plan A and B as follows) for brand i to be available on the platform, or the seller goes to the outside option which yields a payoff of $(\Omega - g(i))$.

i) Plan A. The seller i pays a membership fee r to do business on the platform, but he/she still needs to cover the distribution cost himself/herself. The payoff is thus $D(\mathbf{p})q_i^s(p_i^s - c) - r - g(i)$, where q_i^s is the market share of the brand under Plan A.

ii) Plan B. He/she simply chooses to wholesale his/her brand of products to the platform at price p^w leaving the retailing business to the platform. As he/she neither needs to pay the admission fees nor to incur any distribution costs, the payoff is $D(\mathbf{p})q_i^o(p^w - c)$, where q_i^o is the market share of brand i under Plan B.

Now we have two types of brands. Let us call those under Plan A directly retailed brands and denote their number by n^s; and call those under Plan B platform merchandised brands and denote their number by n^o. Also, we are using superscripts s and o to indicate the directly retailed brands and the platform merchandised brands respectively for the variables, i.e. p_i^s and p_i^o, q_i^s and q_i^o (the subscript i is the index of the brand). It is obvious to note that the price vector $\mathbf{p}$ itself already contains the information of the brand composite, i.e. n^s and n^o.

Timing: We formulate the decisions of the agents in the following sequential game, involving first the platform, then the sellers, and finally the buyers:

Stage 1. The owner of the platform decides to accommodate a total of n brands on the platform, and posts a bidding price p^w for wholesaling as well as a flat membership fee r for all prospective sellers.

Stage 2. The sellers decide whether to pay membership fee to gain access to sell on the platform (as described in Plan A), or to accept the bids from the platform (as Plan B), or to simply go to the outside option. Thus, a market is formed featuring n^s brands directly retailed by independent sellers, together with n^o brands merchandised by the platform.

Stage 3. All the brands that entered in the second stage now enter monopolistic competition. That is, the n^s affiliated direct retailers non-cooperatively decide the prices of their brands and simultaneously the platform decides the prices of its n^o merchandised brands.

Stage 4. By observing all the available brands as well as their prices in the market, the buyers decide which product to buy, if any.

[5] In the formal equilibrium analysis, we assume that ε_i^l is independently and identically double exponentially distributed as commonly seen in Multinomial Logit literature.

2.3 Equilibrium Analysis

Now let us proceed into the equilibrium analysis of the model.

To begin with we look at the decision of the marginal seller. Since the platform posts a single wholesale price p^w, those who go with Plan B will all make same profit, which equals the profit of the marginal seller i^m who is indifferent between taking Plan A and Plan B. That is

$$D(\mathbf{p})q^o_{i^m}(p^w - c) = D(\mathbf{p})q^s_{i^m}(p^s_{i^m} - c) - r - g(i^m) \tag{2.1}$$

$$\Rightarrow \quad r = D(\mathbf{p})(q^s_{i^m}p^s_{i^m} - q^o_{i^m}p^w - q^s_{i^m}c + q^o_{i^m}c) - g(i^m) \tag{2.2}$$

Suppose this marginal seller always chooses to go to Plan A, then we have

$$n^s = i^m$$

Next, it is relatively easy to check that the brands of the same group (i.e. the merchandised brands or the directly retailed brands) follow symmetric pricing strategy in equilibrium, since they all have identical constant marginal cost and face the same market conditions. That is

$$p^s_i = p^s,\ q^s_i = q^s$$

$$p^o_j = p^o,\ q^o_j = q^o$$

where $i = 1, ..., n^s$ and $j = n^s + 1, ..., n^s + n^o$

Now (2.2) can be rewritten as

$$r = D(\mathbf{p})(q^s p^s - q^o p^w - q^s c + q^o c) - g(n^s) \tag{2.3}$$

It is easy to comprehend that in Stage 2 the sellers (or the corresponding brands) with low distribution costs would opt for direct retailing, while those with higher costs would prefer their brands to be merchandised. Recall that once a seller chooses to wholesale his/her products to the platform, it is no longer the seller but the platform that is responsible for the distribution costs. As a result, the brands with index $i = 1, ..., n^s$ are directly retailed by the individual third-party sellers; the brands with index $j = n^s + 1, ..., n^s + n^o$ are merchandised by the platform; and the brands with index higher than $(n^s + n^o)$ are kept away from the marketplace since the platform has committed to host a total of $(n^s + n^o)$ brand variety in Stage 1.

Additionally, the marginal sellers can be identified by observing the information announced in Stage 1. By backward induction, the market share for each brand (i.e. q^s and q^o) is determined by the price $\mathbf{p}$ set in the previous stage. Then in Stage 3, $\mathbf{p}$ (i.e. p^s and p^o but NOT p^w) is uniquely determined by the market participation (i.e. n^s directly retailed brands and n^o platform merchandised brands) formed in the previous stage. Thus, $D(\mathbf{p})$, q^s, q^o and p^s in (2.2) can all be stated in terms of n^s and n^o. We note that the marginal brand i^m (which equals n^s) is determined by n, p^w and r set by the platform in the prior stage[6]. In others words, the brand mix is uniquely determined by the platform's strategy in the first stage.

Next, let us come to the profit contribution for the platform of merchandising brand j. That is

$$\pi_j^o = D(\mathbf{p})q^o(p^o - p^w) - g(j)$$

Thus, the aggregate profit function for the platform can be stated as follows. It consists of two parts: the membership fees collected from n^s subscribed direct retailers, and the profits from the other n^o merchandised brands. That is

$$\Pi = rn^s + \sum_{j=n^s+1}^{n^s+n^o} \pi_j^o \tag{2.4}$$

$$= rn^s + \sum_{j=n^s+1}^{n^s+n^o} [D(\mathbf{p})q^o(p^o - p^w) - g(j)] \tag{2.5}$$

However, two cases need to be discussed separately if additional equilibrium results were to be applied in the platform's profit function. Suppose $\Omega > g(1)$.

Firstly, we look at the case when $n^s \leq g^{-1}(\Omega)$, where $g^{-1}(\Omega)$ is also denoted by n_Ω^s. We note that it is optimal for the platform to set a membership fee that makes the sellers indifferent between choosing Plan A and the outside option. That is

$$D(\mathbf{p})q^s(p^s - c) - r - g(i) = \Omega - g(i) \text{ for all } i \leq n_\Omega^s \tag{2.6}$$

Combining (2.1) and (2.6), we obtain the expressions of p^w and r as follows:

$$r = D(\mathbf{p})q^s(p^s - c) - \Omega \tag{2.7}$$

$$p^w = \frac{\Omega - g(n^s)}{D(\mathbf{p})q^o} + c \tag{2.8}$$

The intuition of (2.7) is that as the value of the outside option Ω increases, the platform will have to lower the membership fee r to remain attractive. The intuition of (2.8) is that in order to "persuade" more sellers to join the direct retailing force, the platform will have to lower the wholesale price p^w to make the merchandising option less attractive. Additionally, (2.8) implies that $p^w \geq c$.

Now, the platform's profit function (2.5) can be manipulated as follows by using (2.7) and

[6] This can be verified by establishing the simultaneous equations of (2.2) and $n^s + n^o = n$.

(2.8).

$$\begin{aligned}
\Pi &= D(\mathbf{p})n^s q^s(p^s - c) - n^s\Omega + \sum_{j=n^s+1}^{n^s+n^o} [D(\mathbf{p})q^o(p^o - p^w) - g(j)] \\
&= D(\mathbf{p})n^s q^s(p^s - c) - n^s\Omega + \sum_{j=n^s+1}^{n^s+n^o} [D(\mathbf{p})q^o(p^o - \frac{\Omega - g(n^s)}{D(\mathbf{p})q^o} - c) - g(j)] \\
&= D(\mathbf{p})(n^s q^s p^s + n^o q^o p^o - c) - (n^s + n^o)\Omega - \sum_{j=n^s+1}^{n^s+n^o} [g(j) - g(n^s)] \qquad (2.9)
\end{aligned}$$

Secondly, we look at the case when $n^s > n^s_\Omega$. It is worth noting that (2.6) and (2.7) will not apply in this case since the net payoff cannot be negative. However, it can be shown in the following lemma that the platform always sets the wholesale price at the marginal cost.

Lemma 2.1 *The platform always sets the wholesale price p^w at the level of marginal cost c when $n^s > n^s_\Omega$.*

Proof. See Appendix. ■

The intuition here is that by setting the wholesale price as low as the marginal cost the platform has effectively "pushed" a larger proportion of seller to join the third-party-brand portfolio. As a result, the platform extracts all the surpluses from the producers of the merchandised brands.

By using the result of Lemma 2.1 that $p^w = c$, (2.3) can be simplified to

$$r = D(\mathbf{p})(q^s p^s - q^s c) - g(n^s) \qquad (2.10)$$

Substituting for r in (2.5) by using (2.10), the platform's profit function can be rewritten as

$$\begin{aligned}
\Pi &= [D(\mathbf{p})(q^s p^s - q^s c) - g(n^s)]n^s + \sum_{j=n^s+1}^{n^s+n^o} [D(\mathbf{p})q^o(p^o - c) - g(j)] \\
&= D(\mathbf{p})(n^s q^s p^s + n^o q^o p^o - c) - n^s g(n^s) - \sum_{j=n^s+1}^{n^s+n^o} g(j) \qquad (2.11)
\end{aligned}$$

Now we have derived two groups of expressions, i.e. (2.4) and (2.9)(2.11), to describe the platform's profits. Both ways bear their own merit.

The profit function (2.4) is straightforward in a sense that each argument represents the source of profits from a certain group of brands.

The profit functions (2.9) and (2.11) can be interpreted in another way without losing the intuition. The first terms in (2.9) and (2.11) are identical, and can be referred to as the "gross profits" (i.e. the profits gross of distribution costs but net of variable costs) generated from both brand types. In order to facilitate our analysis later, let us define $\Psi(n^s, n^o) = D(\mathbf{p})(n^s q^s p^s + n^o q^o p^o - c)$. The last two terms in (2.9) and (2.11) are the "costs", which include the distribution costs plus the profits that go to the direct retailers and the wholesalers. More specifically, the shaded areas under curve $g(i)$ in B in Figure 2.1 represent the distribution costs burdened by the platform for merchandising. The shaded area between curve $g(i)$ and $[g(i) + \Omega - g(n^s)]$ in Figure 2.1(a) is the surplus that goes to the producers of the merchandised brands. The shaded area A represents the profits (gross of distribution costs and membership fees) required by the direct retailers to join the platform rather than going to the outside options. Moreover, it can be interpreted as all the distribution costs incurred (i.e. the shaded area under curve $g(i)$), plus the profit that the platform has to give away to both direct retailer and producers of the merchandised brands (i.e. the shaded area above curve $g(i)$), if any.

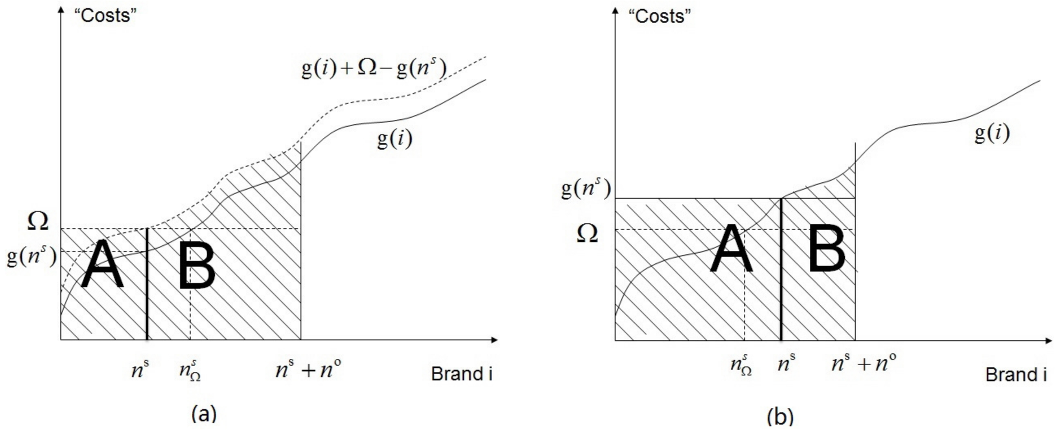

Figure 2.1: (a) "Costs" for the platform when $n^s \leq n^s_\Omega$; (b) when $n^s > n^s_\Omega$

As a matter of fact, the platform decides its optimal level of brand variety n, the wholesale price p^w and the membership fee r in the first stage, taking into account how they will affect competition and profits for all parties in the third and fourth stage and thus the self-selection of sellers in the second stage. Consequently, we can view the platform's planning and pricing problem for admission as its controlling problem over the equilibrium market participation for both directly retailed brands and platform merchandised ones. We establish the following proposition to summarise all these.

Proposition 2.1 *The problem faced by the platform when deciding the total variety of brands n, the wholesale price p^w and the membership fee r can be expressed as a problem of effectively choosing the range of directly retailed brands n^s and the range of platform merchandised brands n^o. Formally, this problem can be written as*

$$\max_{n^s,n^o}\{D(\mathbf{p})(n^s q^s p^s + n^o q^o p^o - c) - (n^s + n^o)\pi^x - \sum_{j=n^s+1}^{n^s+n^o}[g(j) - g(n^s)]\} \text{ when } n^s \leq n^s_\Omega$$

$$\max_{n^s,n^o}\{D(\mathbf{p})(n^s q^s p^s + n^o q^o p^o - c) - n^s g(n^s) - \sum_{j=n^s+1}^{n^s+n^o} g(j)]\} \text{ when } n^s > n^s_\Omega$$

where $D(\cdot)$, p^s and p^o, q^s and q^o are well defined as functions of n^s and n^o.

Let us go back and explore a bit further the mass of the consumers who eventually make purchases. As we mentioned before, a buyer purchases a brand only if net utility derived exceeds that of all other brands and is non-negative. Given the vector of prices $\mathbf{p}$ set in the previous stage, the expected per capita demand for brand i can be written as follows:

$$x_i(\mathbf{p}, n) = \Pr[v_i^l \geq \max\{\max_{j \neq i}\{v_j^l\}, 0\}]. \tag{2.12}$$

Recall that the indirect utility function for a typical buyer l purchasing brand i is $v_i^l = v - p_i + \varepsilon_i^l$. We assume that ε_i^l, the stochastic term of consumer l's utility function, is independently and identically distributed according to a type-1 Extreme Value (EV) distribution with location parameter zero and scale parameter μ.[7] Here, μ can be interpreted as a parameter of substitutability among the brands. Using the EV distribution is appealing because it is closely related to the familiar logit model, which can be conveniently modified to allow for a no-purchase alternative. Another rationale for using EV distribution as a preference parameter is "to keep the demand structure analytically tractable" (Jain, et. al., 1994). Also, Dolan (1995) gives empirical evidence that the reservation prices (which is the preference in our model) for different but related items often have symmetric distributions with the same shape. While the extreme value distribution is not exactly symmetric, it can be viewed as an approximation to the Normal distribution.

Given the indirect utility function of consumers and demand function (2.12), the per capita demand for brand i can be written as follows[8].

$$x_i(\mathbf{p}, n) = [\frac{e^{(v-p_i)/\mu}}{M}][1 - e^{-M}] \tag{2.13}$$

$$\text{where } M = \sum_{j=1}^{n} e^{(v-p_j)/\mu} = e^{v/\mu} \sum_{j=1}^{n} e^{-p_j/\mu}$$

We need to note that the first term in the square brackets of (2.13) is the market share for brand i,

$$q_i(\mathbf{p}) = \frac{e^{(v-p_i)/\mu}}{M} = \frac{e^{-p_i/\mu}}{\sum_{j=1}^{n} e^{-p_j/\mu}} \tag{2.14}$$

[7] The $EV(0, \mu)$ distribution is unimodal with mode zero and is skewed to the right. Its cumulative distribution and density functions are given by

$$\begin{aligned} F(x) &= \exp\{-e^{-x/\mu}\} \\ f(x) &= \mu^{-1} e^{-x/\mu} F(x) \end{aligned}$$

[8] Based on the indirect utility function we've stated earlier for the consumers and the monopolistic competition among brands, we are actually using the Multinomial Logit Demand.

We write q_i for $q_i(\mathbf{p})$ in the following paragraphs for the sake of brevity, but keep in mind that market share is shaped by market prices $\mathbf{p}$ which is determined by (n^s, n^o).

The second term in the square brackets of (2.13) indicates the market demand $D(\mathbf{p})$, which is the mass of consumers that finally make successful purchases. It is worthwhile to remember two important derivative properties of the market share as follows.

$$\frac{\partial q_i}{\partial p_j} = \frac{q_i q_j}{\mu} > 0 \tag{2.15}$$

$$\frac{\partial q_i}{\partial p_i} = \frac{q_i(q_i - 1)}{\mu} < 0 \tag{2.16}$$

Next, we characterise the pricing strategy for each party in Stage 3.

The profit function for the direct retailer with brand i is

$$\pi_i^s = D(\mathbf{p})(p^s - c)q^s - r - g(i)$$

The best response function for the direct retailer can be obtained by differentiating the above function with respect to p^s. Suppose that the individual brands ignore their impact on the market demand $D(\mathbf{p})$, as supported by Chamberlin's discussion of "large group" equilibrium. Also, since r and $g(i)$ have already been incurred in the previous stage, we are effectively searching for the derivative of $(p^s - c)q^s$. Thus, the first order condition can be obtained as follows by using (2.15) and (2.16).

$$\begin{aligned} & q^s + (p^s - c)\frac{\partial q^s}{\partial p^s} \\ = \quad & q^s + \frac{(p^s - c)q^s(q^s - 1)}{\mu} = 0 \end{aligned} \tag{2.17}$$

Rearranging the above, we obtain an implicit best response function for the direct retailers. That is

$$p^s - c = \frac{\mu}{1 - q^s} \tag{2.18}$$

Next, let us turn to the pricing strategy for platform merchandised brands. Similarly, we are effectively searching for the derivative of $(n^s q^s p^s + n^o q^o p^o)$. However, we need to note that the best response for the merchandised brands cannot be obtained by simply differentiating it with respect to p^o[9]. Instead we need to separate a brand j from the n^o merchandised brands

[9] Differentiating $(n^s q^s p^s + n^o q^o p^o)$ with respect to p^o would indicate simultaneous price change of *all* merchandised brands. In this case, we cannot ignore their joint impact on the market demand $D(\mathbf{p})$ and the "large group" assumption breaks down.

and differentiate the platform's profit function with respect to p_j^o. That is

$$
\begin{aligned}
& n^s p^s \frac{\partial q^s}{\partial p_j^o} + q_j^o + p_j^o \frac{\partial q_j^o}{\partial p_j^o} + (n^o - 1) p_i^o \frac{\partial q_i^o}{\partial p_j^o} \\
= \quad & \frac{n^s p^s q^s q_j^o}{\mu} + q_j^o + \frac{p_j^o q_j^o (q_j^o - 1)}{\mu} + \frac{(n^o - 1) p_i^o q_i^o q_j^o}{\mu} = 0 \qquad (2.19)
\end{aligned}
$$

$$\text{where } i \neq j$$

In equilibrium all the prices (market shares) of the merchandised brands shall be same. After a bit more simplification, we obtain a neat implicit best response function for the platform merchandised brands:

$$p^o - p^s = \frac{\mu}{1 - n^o q^o} = \frac{\mu}{n^s q^s} \qquad (2.20)$$

Since $n^o q^o$ is strictly less than unity in the context of a mixed brand portfolio, it is straightforward to have the following corollary.

Corollary 2.1 *The platform merchandised brands are always priced higher than the directly retailed brands; that is,* $p^o > p^s$.

The intuition of Corollary 2.1 is that, on the one hand, the platform has no incentive to undercut the directly retailed brands, as it would lower the profit margin of the "seemingly rival" brands which contribute to the platform's profit in terms of the membership fees. On the other hand, the platform merchandised brands face less competition than the directly retailed brands since the pricing strategy of the former is under the control of a single ownership (i.e. the platform)[10].

Alternatively, Corollary 2.1 can also be understood in the following way:

We learn from the difference between the reaction functions (2.19) and (2.17) that the (upward-sloping) best response curve of the platform merchandised brand is the rightward shifted mirror image of the best response curve of the directly retailed brand (as illustrated in Figure 2.2). It can be easily implied from the figure that the platform merchandised brands always charge a price premium.

2.3.1 Existence of the Optimal Brand Portfolio

In this subsection, we examine the existence of the optimal brand portfolio with a fixed level of brand variety, i.e. $n^s + n^o = \bar{n}$. The virtue of adopting a fixed brand variety is obvious: on the one hand, it helps to separate the indirect network effect in consumer demand which would complicate the analysis from the very beginning; on the other hand, it allows us to scrutinise

[10] This intuition of price premium here is in line with the concept of "price shield" suggested by Giraud-Heraud et. al. (2003) in the context of competition among a multi-product firm and many single-product firms.

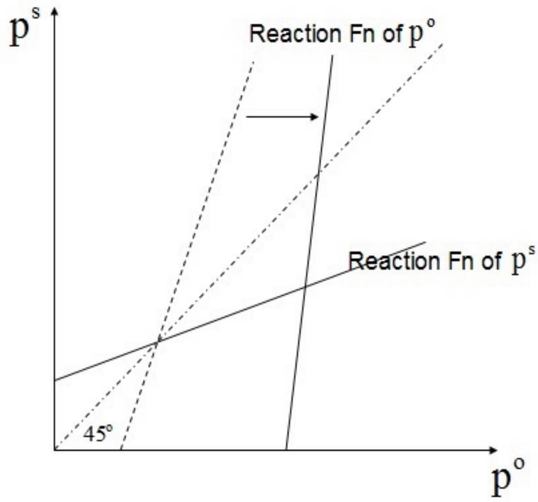

Figure 2.2: Illustration of the equilibrium prices of the two brand types

the platform's incentive for whether or not to absorb one more brand to merchandise. Also in the following analysis, we ignore the integer constraints and treat the number of the directly retailed brands n^s as continuous variables, where $n^s \in 0 \cup [1, \bar{n}]$;[11] and treat the distribution cost $g(i)$ as a continuous function, where $i \in [1, \bar{n}]$.

If the number of brands are in hundreds or thousands, the continuous function of the distribution cost is a good proxy of the discrete problem. Thus, the results derived in the previous subsection still hold. Recall from Corollary 2.1 that the platform merchandised brands charge a higher price than the directly retailed ones. It implies the latter enjoys a larger market share than the former. Then, if an existing directly retailed brand were to be absorbed by the platform (i.e. n^s decreases while n^o increases), we would anticipate its price to be inflated and market share lowered. Meanwhile, it would leave more market share to its rivals and allow them also to raise their prices a little bit. The overall effect is that all prices would be inflated (but the prices of the counterparts would not rise as much) and the per customer "gross profits" enhanced as the platform absorbs more brands to merchandise while keeping the total brand variety fixed. However, the mass of the consumers shrinks as a result of being discouraged by the price inflation. We establish the following lemma to summarise the movements of the relevant market variables.

Lemma 2.2 *With a fixed brand variety on the platform (i.e. $n^s + n^o = \bar{n}$),*

(i) the prices charged for both the platform merchandised brands (i.e. p^o) and the directly retailed brands (i.e. p^s), and the price premium for the merchandised brands (i.e. $p^o - p^s$)

[11] It is important to note that $n^s \notin (0, 1)$, otherwise the reaction functions will be undefined for the market competitors in the Multinomial Logit Setups.

decrease in n^s*;*

(ii) the market share for each platform merchandised brand (i.e. q^o*) and for each directly retailed brand (i.e.* q^s*)*[12] *decrease in* n^s*;*

(iii) the market share for platform merchandising (direct retailing), i.e. $n^o q^o$ $(n^s q^s)$*, decreases (increases) in* n^s*;*

(iv) the market demand $D(n^s, \bar{n} - n^s)$ *increases in* n^s*;*

(v) the "gross profits" per consumer (i.e. $n^s q^s p^s + n^o q^o p^o - c$*) decreases in* n^s*;*

(vi) the "costs" (i.e. $(n^s + n^o)\Omega + \int_{n^s}^{\bar{n}} [g(j) - g(n^s)]dj$*) decrease in* n^s *when* $n^s \leq n^s_\Omega$*; while the "costs" (i.e.* $n^s g(n^s) + \int_{n^s}^{\bar{n}} g(i)di$*) increase in* n^s *when* $n^s > n^s_\Omega$*.*

(NB: All the above relations with respect to n^o *are just the opposite since* $n^o = \bar{n} - n^s$*)*

Proof. By using (2.14), equation (2.20) can be expanded as

$$p^o - p^s = \frac{\mu}{1 - n^o q^o} = \mu + \frac{(\bar{n} - n^s)\mu}{n^s} \exp(\frac{p^s - p^o}{\mu}) \tag{2.21}$$

By using the Implicit Function Theorem, we obtain

$$\frac{d(p^o - p^s)}{dn^s} = \frac{-\mu\bar{n}\exp(\frac{p^s - p^o}{\mu})}{n^{s2} + n^s(\bar{n} - n^s)\exp(\frac{p^s - p^o}{\mu})} \tag{2.22}$$

Since $\exp(\frac{p^s - p^o}{\mu}) > 0$, we can conclude that $\frac{d(p^o - p^s)}{dn^s} < 0$.

By using (2.20) and monotonicity, it can be further implied that $\frac{dq^o}{dn^s} < 0$.

Additionally, it can be implied from (2.14) that $\frac{q^s}{q^o} = \exp(\frac{p^o - p^s}{\mu})$. By using the subconclusions above, it implies $\frac{dq^s}{dn^s} < 0$.

Then, we arrive at $\frac{\partial p^s}{\partial n^s} < 0$ by using (2.18); as well as $\frac{\partial p^o}{\partial n^s} < 0$ by using (2.20). So far, we have proved (i) and (ii) in the lemma.

It is also straightforward to arrive at (iii) from (ii).

Next, let us turn to $D(n^s, \bar{n} - n^s)$, the mass of the consumers who eventually make purchases. We have

$$\begin{aligned} D(n^s, \bar{n} - n^s) &= 1 - e^{-M} = 1 - \exp[-e^{v(\bar{n})/\mu}(n^o e^{-p^o/\mu} + n^s e^{-p^s/\mu})] \\ &= 1 - \exp[\frac{-e^{v(\bar{n})/\mu} e^{-p^o/\mu}}{q^o}] \end{aligned}$$

Recall that $\frac{\partial p^o}{\partial n^s} < 0$ in Lemma 2.2 (i) and $\frac{\partial q^o}{\partial n^s} < 0$ in (ii). By using monotonicity it can be shown that $D(n^s, \bar{n} - n^s)$ increases in n^s.

For (v), by using (2.18) and (2.20), the "gross profits" per consumer can be simplified as

---

[12] Here q^o and q^s are not infinitesimally small if n is finite.

follows.

$$n^s q^s p^s + n^o q^o p^o - c = p^o - \mu - c \tag{2.23}$$

Since p^o decreases in n^s (Lemma 2.2 (i)), it is straightforward that the "gross profits" (i.e. $n^s q^s p^s + n^o q^o p^o - c$) decreases in n^s.

For the last statement in the lemma, we can write as follows the derivative of the "costs" with respect to n^s when $n^s \leq n^s_\Omega$.

$$\frac{d(\bar{n}\Omega + \int_{n^s}^{\bar{n}} [g(j) - g(n^s)]dj)}{dn^s} = -(\bar{n} - n^s)g'(n^s) < 0 \tag{2.24}$$

The negative sign of the above derivative indicates that when $n^s \leq n^s_\Omega$ the "costs" decrease in n^s. It can also be implied from the change in the shaded area in Figure 2.1 (a).

Similarly, the derivative of the "costs" with respect to n^s when $n^s > n^s_\Omega$ can be written as

$$\frac{d[n^s g(n^s) + \int_{n^s}^{\bar{n}} g(j)dj]}{dn^s} = n^s g'(n^s) > 0 \tag{2.25}$$

The positive sign of the above derivative indicates that when $n^s > n^s_\Omega$ the "costs" increase in n^s. It can also be implied from the change in the shaded area in Figure 2.1 (b).

Here we have completed all the proofs in Lemma 2.2. ■

The intuition of Lemma 2.2 is relatively easy to grasp. Since its profits come from merchandising and renting, the platform can more credibly commit to putting up higher prices of its merchandised brands by holding a larger stake in the retailing market, which justifies its incentive to merchandise more brands. Moreover, due to the complementarity nature of prices, the directly retailed brands are priced higher, while their counterparts (i.e. the merchandised brands) are priced even higher as the platform enriches its self-brand portfolio. Additionally, we find the coexistence of the widened price discrepancy between the two brand groups and an increase in market share for each brand, which is unusual and might be of interest to discuss more.

From the "gross profits" side, Lemma 2.2 (v) justifies the platform's incentive to expand the range of the merchandised brands. However, we also need to note that the expansion in the platform's self-brand portfolio pushes up the market prices and will in turn dampen the consumers' motivation to visit the platform in the first place (Lemma 2.2 (iv)). Therefore, it implies an optimal mix of platform merchandised brands and directly retailed brands for the platform.

The last statement in Lemma 2.2 captures the intuition on the "costs" side for the platform (see the illustration in Figure 2.1). It is easy to check from Figure 2.1 (a) and (b) that the platform tends to keep the "costs" low at $n^s = n^s_\Omega$.

A formal claim on the existence as well as the type of equilibrium will be presented in Proposition 2.2. Before that, we need to establish Corollary 2.2 to have a more detailed characterisation of the "gross profits" side function for the platform.

Corollary 2.2 $\frac{d\Psi(n^s, \bar{n}-n^s)}{dn^s}|_{n^s=1} > 0$ *and* $\frac{d^2}{dn^{s2}}\Psi(n^s, \bar{n}-n^s) < 0$.

Proof. See Appendix. ■

Corollary 2.2 simply states that Ψ is concave and the slope of Ψ function at $n^s = 1$ is always positive. It is worth noting that the sufficient condition to yield a unique equilibrium is that the upward-sloping proportion of the "costs" curve (i.e. $n^s g(n^s) + \int_{n^s}^{\bar{n}} g(i)di$) is convex.

Recall that we are now discussing the case where $n^s \in [1, \bar{n}]$; it helps to have the following definition by using the derivative of upward-sloping part of the "costs" function (2.25).

Definition 2.1 *The threshold scale of the brand portfolio* $\bar{n}_0$ *is implicitly defined by*

$$\frac{\partial\Psi(n^s, \bar{n}_0 - n^s)}{\partial n^s}|_{n^s=\bar{n}_0} = \bar{n}_0 g'(\bar{n}_0) \tag{2.26}$$

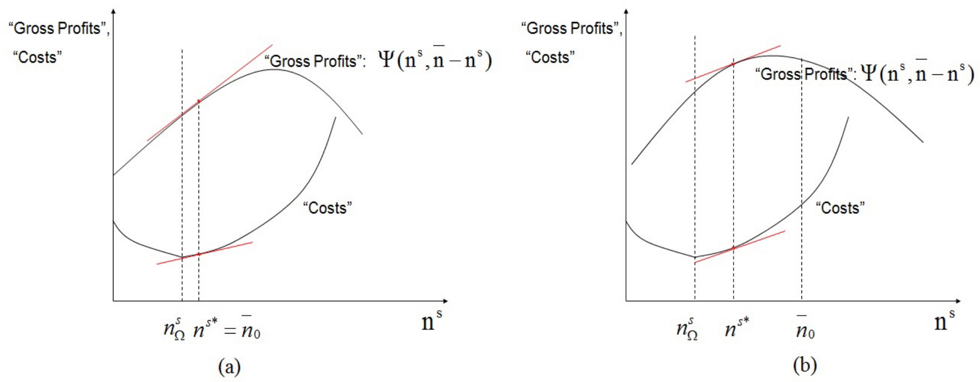

Figure 2.3: (a)The "gross profits" and "costs" for the platform when $\bar{n}$ is relatively small; (b)when $\bar{n}$ is relatively large.

From Figure 2.3 (a) it is not difficult to justify the existence of a corner solution at $n^s = \bar{n}_0$ when $\bar{n} \leq \bar{n}_0$; and from Figure 2.3 (b) the existence of an interior solution when $\bar{n} > \bar{n}_0$. This can be summarised in the following proposition.

Proposition 2.2 *(i) If* $\bar{n} \leq \bar{n}_0$ *(i.e. when the platform's capacity is relatively small), only the directly retailed brands dominate the market in equilibrium;*

(ii) If $\bar{n} > \bar{n}_0$ *(i.e. when the platform's capacity is relatively large), both the platform merchandised brands and the directly retailed brands coexist in the marketplace.*

Proposition 2.2 is the main result of this subsection. It implies that when a platform's scale is small it would favour a pure two-sided business model, since the marginal "gross profits" brought by a directly retailed brand rather than a merchandised brand exceeds its marginal "costs"; and when a platform's scale is large, it would be tempted to replace some directly retailed brands with merchandised ones, since now it is able to share the high profit margins by merchandising, leaving less surplus to the direct retailers.

We note that the interior solution n^{s*} is always greater than n^s_Ω. Thus, we learn from Lemma 2.1 that, the platform always sets the wholesale price at the marginal cost and extracts all the surplus from the producers of the merchandised brands in the hybrid business model.

The following numerical example helps us to have a glimpse of the equilibrium market profile of the two cases in Proposition 2.2.

Example 2.1 *Given that the total variety of brands on the platform is fixed at (a)* $n^s + n^o = \bar{n} = 4$ *and (b)* $n^s + n^o = \bar{n} = 10$*, the constant marginal cost* $c = 0.1$*, the parameter of brand substitutability* $\mu = 0.1$*, the value of the outside option* $\Omega = 3/5000$*, the systematic value of the brands* $v = \frac{10^{1/2}}{50}$*, and the distribution cost* $g(x) = x/5000$*. Then the optimal portfolio for the platform and the relevant market profile is displayed in Table 2.1.*

	$\bar{n}$	n^o	n^s	p^o	p^s	q^o	q^s	D	Π^{o*}
Case (a)	4	0	4	n/a	0.233	n/a	0.250	0.839	0.06909
Case (b)	10	5.70	4.30	0.355	0.221	0.045	0.173	0.698	0.09656

Table 2.1: The optimal brand portfolios when $\bar{n} = 4$ and when $\bar{n} = 10$

We note that $n^s_\Omega = g^{-1}(\Omega) = g^{-1}(3/5000) = 3$.

Moreover, recall that $\bar{n}_0$ is defined by equation (2.26), thus, the threshold level of capacity for the platform in this example can be derived by using (2.33) in the Appendix. That is

$$
\begin{aligned}
& \frac{\partial \Psi(n^s, \bar{n}_0 - n^s)}{\partial n^s}\Big|_{n^s=\bar{n}_0} \\
= \quad & -e^{-\bar{n}_0 e^{-\frac{1}{1-1/\bar{n}_0} - \frac{0.1}{0.1} + \frac{10^{1/2}}{5}}} e^{\frac{10^{1/2}}{5}} [e^{-\frac{1}{1-1/\bar{n}_0} - 1 - \frac{0.1}{0.1}} - e^{-\frac{1}{1-1/\bar{n}_0} - 1} \\
& - \frac{\bar{n}_0}{0.1} e^{-\frac{1}{1-1/\bar{n}_0} - 1} \cdot \frac{(1-e^{-1})0.1}{(\bar{n}_0 - 1)^2}] \frac{0.1}{1 - 1/\bar{n}_0} \\
& -(1 - e^{-\bar{n}_0 e^{-\frac{1}{1-1/\bar{n}_0} - \frac{0.1}{0.1} + \frac{10^{1/2}}{5}}} e^{\frac{10^{1/2}}{5}})(\frac{e^{-1}0.1}{\bar{n}_0} + \frac{(1-e^{-1})0.1}{(\bar{n}_0 - 1)^2}) \\
= \quad & \bar{n}_0/5000 \\
\Rightarrow \quad & \bar{n}_0 = 4.71
\end{aligned}
$$

It can be verified that the results in Table 2.1 coincide with the threshold levels of capacity derived here.

2.3.2 Comparative Statics w.r.t. the Hosting Capacity

In this subsection, we focus on the comparative statics with respect to the total brand variety under Proposition 2.2 (ii) (i.e. $\bar{n} > \bar{n}_0$). It helps us to gain some intuition about the dynamics of optimal brand portfolio when the platform's hosting capacity $\bar{n}$ expands. Our analysis shows a backward bending proportion plus a vertical proportion of the "contract curve"; that is, when the platform's accommodating capacity $\bar{n}$ expands from $\bar{n}_0$, the optimum number of the platform merchandised brands n^o always increases while that of the directly retailed brands n^s decreases until it reaches n^s_Ω.

Recall the best response functions of the two brand types (2.18) and (2.20), which can be expanded as

$$p^o - p^s = \frac{\mu n^o}{n^s} \exp[-(p^o - p^s)/\mu] + \mu \tag{2.27}$$

$$p^s - c = \mu/[1 - \frac{1}{n^o \exp[-(p^o - p^s)/\mu] + n^s}] \tag{2.28}$$

Here is a system of two equations with four unknowns: p^o and p^s, n^o and n^s. We could solve for p^o and p^s in terms of n^o and n^s[13]. And the optimal portfolio of the brands could be identified by solving the above two-dimensional optimisation as described in Proposition 2.1. However, the problem lies in the fact that there is no closed-form solution for $(p^o - p^s)$ in (2.27).

We concluded in the previous subsection that there exists a unique equilibrium for every level of brand variety (e.g. $E_{n^s+n^o}$ and $E_{n^s+n^o+1}$ on the -45^o lines). Connecting the Es we obtain a "contract curve" (see the illustration of the curve in Figure 2.4). Thus, identifying the optimal brand portfolio in the hybrid mode case is no more than finding the brand portfolio along the "contract curve" that gives the highest profits for the platform. However, the direct analysis of the change from $E_{n^s+n^o}$ to $E_{n^s+n^o+1}$, for instance, would be too complex due to the endogeneity of such intermediary variables as p^o and p^s which can not be explicitly expressed in terms of the choice variables n^o and n^s. Nonetheless, in the following paragraphs we provide a novel way to decompose the effect of the change from $E_{n^s+n^o}$ to $E_{n^s+n^o+1}$.

Firstly, we establish Corollary 2.3 to answer a simpler question as to how Ψ, the "gross profits" for the platform, responds to the platform's shift of its portfolio from $E_{n^s+n^o}$ to $E'_{n^s+n^o+1}$, i.e. the platform fills in a newly added slot with platform merchandised brand holding the number of the directly retailed brands constant[14]. Secondly, we consider the impact on the "gross profits" for the platform Ψ from portfolio $E'_{n^s+n^o+1}$ to $E_{n^s+n^o+1}$ which is the optimal portfolio when the platform's scale is $(n^s + n^o + 1)$.

We show in Corollary 2.3 that $\Psi(n^s, n^o + 1) > \Psi(n^s, n^o)$. That is, Ψ curve shifts upward as the platform's accommodating capacity expands (see Figure 2.5).

Corollary 2.3 *$\Psi(n^s, n^o)$ increases in n^o, holding n^s constant; that is, with a fixed variety of the directly retailed brands, the "gross profits" increase in the variety of the platform merchandised brands.*

[13] We can first solve for $(p^o - p^s)$ in terms of n^o and n^s from equation (2.27), and then substitute it into (2.28) to solve for p^s. Thus, p^o and p^s are both solved in terms of n^o and n^s.

[14] In order to simplify the analysis, here we treat the brand numbers as discrete.

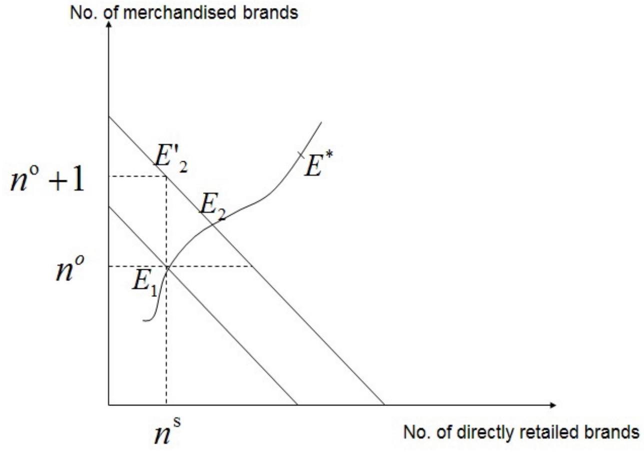

Figure 2.4: Illustration of the optimal brand mix at each fixed level of brand variety

Proof. See Appendix. ■

It is worthwhile to note the difference between the statements in Corollary 2.2 and Corollary 2.3. The former describes how Ψ responds to the change in the composite of market brands by holding the *total* variety constant, while the latter describes the case by holding the variety of the *directly retailed* brands constant.

We show in Corollary 2.4 that, when the platform's accommodating capacity expands, the slope of the tangent line for Ψ curve becomes smaller (see Figure 2.5).

Corollary 2.4 $\frac{\partial^2}{\partial n^s \partial \bar{n}} \Psi(n^s, \bar{n} - n^s) < 0.$

Proof. See Appendix. ■

Corollary 2.4 states that the sensitivity of the portfolio composite to the "gross profits" at every level is mitigated/diluted by the expanding portfolio scale.

Meanwhile, let us turn to the "costs" side of the platform. We learn from (2.24) that when $n^s \leq n^s_\Omega$ the derivative of the "costs" decreases in the platform's scale $\bar{n}$ and from (2.25) that when $n^s > n^s_\Omega$ the "costs" are independent of $\bar{n}$. It can be implied from Figure 2.5 that when $\bar{n}$ increases the optimum number of the directly retailed brands n^{s*} decreases until it reaches n^s_Ω. We summarise it in the following proposition.

Proposition 2.3 *As the platform's accommodating capacity $\bar{n}$ expands, the optimum number of the platform merchandised brands n^o always increases, while the directly retailed brands n^s decreases until it reaches n^s_Ω.*

Proposition 2.3 states that the "contract curve" has a backward bending proportion and a vertical proportion. Combining Proposition 2.2 and Proposition 2.3, we can illustrate the optimal brand portfolio for the platform for every level of brand variety $\bar{n}$ (see the thick line in Figure 2.6).

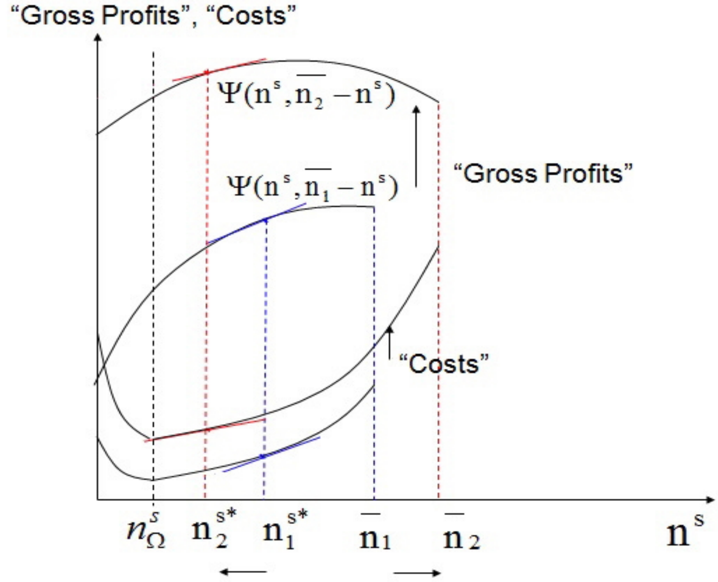

Figure 2.5: The changes in the shape of the "gross profits" and "costs" curves when the total brand variety increases

At first glance the backward bending "contract curve" seems to be counter-intuitive, as many would anticipate the increase in both brand types when the platform was able to host more brands. However, we need to note that as the marginal contribution of the directly retailed brand on "gross profits" is diluted by the expanding capacity, the platform would choose to accommodate more merchandised brands rather than directly retailed brands as it capacity increases.

Note that the discussion in this subsection so far is based on the assumption that $n_{\Omega}^{s} < \bar{n}_0$. It is relatively easy to comprehend that the backward bending proportion of the "contract curve" degenerates to the vertical proportion when $n_{\Omega}^{s} \geq \bar{n}_0$ (See the illustration in Figure 2.7).

The insight derived here is also supported by the empirical evidence. As platforms normally use capacity expansion as one of the most important strategies,[15] for those that accommodate both brand types as we have discussed, many have witnessed an increasing weight and focus on the merchandise (or the self-brands) in the brand portfolio. Among them, the most prominent example is Amazon.com. Started as an online bookstore, Amazon.com has now expanded its product lines into almost anything you can come up with. Meanwhile, it also reaps about 40 percent of its sales from the affiliated direct retailers who do business through "Amazon Marketplace" by paying commission fees. As its accommodating capacity expands it is also interesting to observe that, by offering warehouse space and/or logistic service at a relatively

[15] This strategy is more common for the virtual marketplaces online as they can easily expand the hosting capacities by investing in more servers; and for the traditional business such as supermarket, it is able to increase the capacity by more efficient use of its shelves.

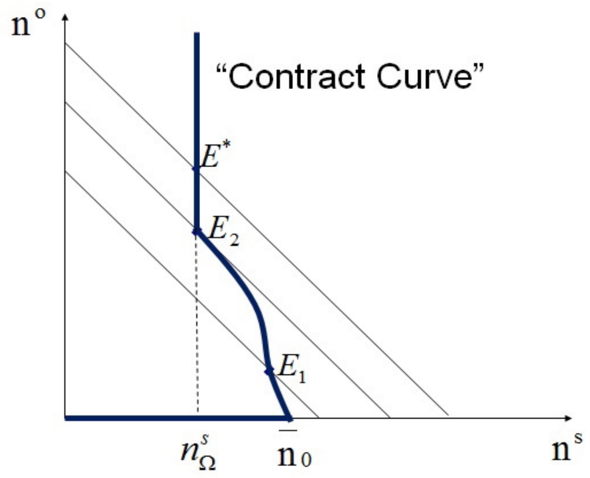

Figure 2.6: Illustration of the "contract curve"

low cost to individual retailers, Amazon has in effect integrated these affiliated brands into its merchandised brand portfolio. Amazon is not alone. Another Chinese online shopping portal, TaoBao Marketplace is now refocusing its core business from its C2C merchants to B2C Taobao Mall merchants. Our model sheds some light on the motivation beneath this trend, though it can be attributed to other factors as well.

2.3.3 Discussion of Indirect Network Effect

Notice that our discussion so far has been restricted to the degenerated form of the systematic value which is v. We are also interested in the impact of positive indirect network effect in consumer demand (i.e. $v = v(\bar{n})$ which increases in $\bar{n}$)[16] on the optimal portfolio for every level of brand variety. In this subsection, we establish the following lemma to specify how our previous results are affected by the introduction of indirect network effect in consumer demand. This will be followed by a numerical example which compares the optimal brand portfolios with and without the indirect network effect.

Lemma 2.3 *When there is indirect network effect in consumer demand and $M > 1$, the optimal portfolio will have more platform merchandised brands for every level of brand variety.*

Proof. See Appendix. ■

We note that consumer demand is a monotonic transformation of M, which is $D(n^s, \bar{n}-n^s) = (1 - e^{-M})$. Thus, M can be viewed as an index of market demand. Also, recall the expression of M, which is $M = (\bar{n} - n^s)e^{(v(\bar{n})-p^o)/\mu} + n^s e^{(v(\bar{n})-p^s)/\mu}$ where $n^s \in [1, \bar{n}]$ and $\bar{n} \geq 2$. It can be

[16] Or it can be attributed to the shorten "psychological distance" to the desirable brands when buyers have more options available.

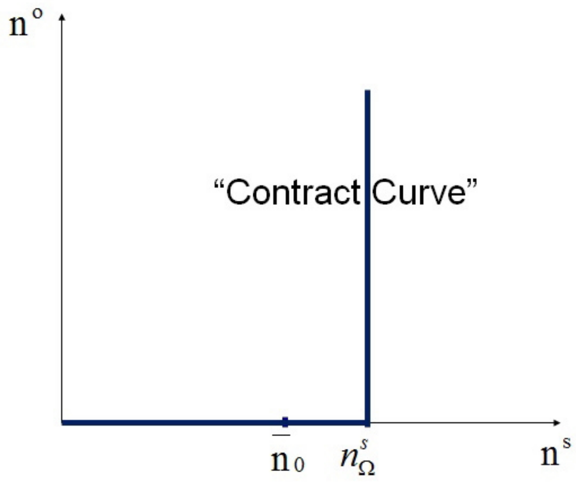

Figure 2.7: Illustration of the "contract curve" when $n^s_\Omega \geq \bar{n}_0$.

shown that M is strictly greater than unity when $v(\bar{n}) > p^o$ and $v(\bar{n}) > p^s$. We note that these two inequalities are true in the cases that the *expected* surplus for any consumer is positive.

Here, the effects are two fold: firstly, the market demand has been reinforced by the indirect network effect as consumers welcome more variety of brands; secondly, when the consumer demand is above a certain level (i.e. the index of market demand M exceeds unity), despite the greater distribution costs to be incurred the platform is willing to merchandise more brands, and hence is able to extract more surplus from the direct retailers (as Area A shrinks in Figure 2.1 (b)). As a result, with additional indirect network effect the platform is better off while the direct retailers worse off. The effect on the optimal brand portfolio is displayed in Figure 2.8. That is, the threshold level of scale $\bar{n}_0$ becomes smaller and the backward bending proportion of the "contract curve" shifts leftward with the vertical proportion remaining still.

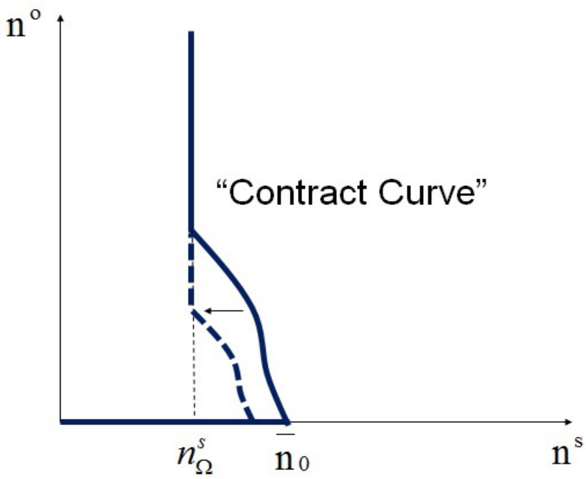

Figure 2.8: The impact of indirect network effect on the "contract curve" when $M > 1$

Additionally, without the restriction on total brand variety as in Example 2.1 (i.e. $n^o + n^s = \bar{n} = 10$), we are able to numerically compare the optimal brand portfolio *without* indirect network effect (i.e. $v(n) = v = \frac{10^{1/2}}{50}$) and that *with* indirect network effect (i.e. $v(n) = \frac{n^{1/2}}{50}$) in Table 2.2.

	$\bar{n}$	n^o	n^s	p^o	p^s	q^o	q^s	D	Π
Optimal brand mix w/o indirect network effect	16.82	12.8	4.02	0.381	0.218	0.030	0.153	0.751	0.10578
Optimal brand mix with indirect network effect	21.82	18.8	3.02	0.413	0.221	0.025	0.173	0.802	0.12183

Table 2.2: The comparison of the optimal brand portfolios with/without indirect network effect

It is not surprising to see from the above numerical example that neglecting the indirect network effect will create a bias toward a smaller number of total brand varieties.

Recall that $n^s_\Omega = 3$, which is smaller than the number of directly retailed brands in the optimal brand mix. Thus, the optimal brand portfolios in both cases are located on the backward bending proportion of the "contract curve".

We learn from Table 2.2 that when we take the indirect network effect into account, the self-brand portfolio expands, while the third-party-brand portfolio shrinks. Additionally, we can also judge this issue in terms of the units sold rather than the variety available for each portfolio type. We note that, as the per brand market share for directly retailed brands increases while that for the platform-merchandised brands decreases[17], the gap in units sold turns out to be not as large as that in brand variety.

2.3.4 Discussion of Outside Option

We note that n^s_Ω is crucially dependent on the value of the outside option (i.e. $n^s_\Omega = g^{-1}(\Omega)$). It can be implied that the backward bending proportion of the "contract curve" prolongs and the vertical proportion shifts leftward with $\bar{n}_0$ unchanged when the outside option becomes less attractive (see the illustration in Figure 2.9). And intuitively, the platform is able to extract more surplus from the direct retailers[18]. In the extreme case where the value of the outside option is not even attractive (i.e. $\Omega \leq g(i)$ where $i = 1, ..., n$), the platform prefers to be a pure "merchant" at each level of brand variety (i.e. the "contract curve" becomes a vertical line at $n^s = 0$).

[17] The reason is that the price of the directly retailed brands increases less than that of the merchandised brands.

[18] The platform has already extracted all the surpluses from the producers of the merchandised brands by setting the wholesale price at the marginal cost.

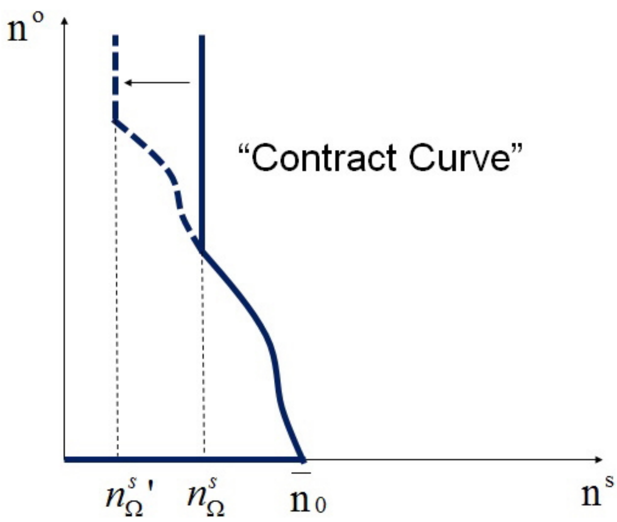

Figure 2.9: Illustration of the "contract curve" when the outside option becomes less attractive

2.4 Conclusion

This chapter has derived two key insights concerning the optimal composite of the two brand types for a two-sided platform that connects consumers and brands sold by both direct retailers and those merchandised by the platform.

Firstly, we examine the existence of the optimal brand portfolio for every level of brand variety. We find that the intermediary would prefer to be a pure "two-sided platform" when the platform's hosting capacity is below a certain level; and both brand types coexist on the platform when its hosting capacity is above a certain level. We have mainly discussed the hybrid mode and the trade-offs between the "intermediary-as-a-middleman" and the "intermediary-as-a-two-sided-platform" business models from the perspective of the platform owner. Our analysis shows that the merchandised brands are always priced higher than the directly retailed ones, and the platform extracts all the surpluses from the producers of the merchandised brands in the hybrid mode. We also find that whenever the platform includes an existing directly retailed brand into its self-brand portfolio, the equilibrium prices of both types of brands are increased.

Secondly, the comparative statics with respect to the platform's hosting capacity capture some intuition about the dynamics of the optimal brand portfolio. We find a backward bending proportion plus a vertical proportion of the "contract curve", which means that when the intermediary increases its capacity, the self-brand portfolio expands while the third-party-brand portfolio shrinks until it reaches a certain level. This result is supported by empirical evidence such as the history of business expansion for Amazon.com. Additionally we show that with indirect network effect in consumer demand and positive expected surplus for the consumers, the platform is able to extract more surplus from the direct retailers. Consequently, the platform is better off while the direct retailers are worse off. Also, we find that the reduction in the value of the outside option will allow the platform to extract more surplus from the direct retailers, and the pure "merchant" mode is preferred only when the outside option is not at all attractive.

There are a few possible extensions of our analysis that are of interest for future research. One would be a more general form of the competition among market participants (especially not restricted by the discreet feature of the Multinomial Logit demand structure). Another would be the consideration of the heterogeneity even within the same type of brands, which would bring the issue of asymmetric pricing (within the same brand type), e.g. the location problem of the Salop's circular model other than the spaceless MNL model in this chapter.

2.5 Appendix

Proof of Lemma 2.1. We need to go back to the platform's profit function (2.5) which can be expanded by using (2.3). That is

$$\Pi = D(\mathbf{p})n^s(q^sp^s - q^op^w - q^sc + q^oc) - n^sg(n^s) + \sum_{j=n^s+1}^{n^s+n^o}[D(\mathbf{p})q^o(p^o - p^w) - g(j)] \qquad (2.29)$$

Recall that all the elements of the market profile in the above expression (i.e. $D(\mathbf{p})$, q^s, q^o, p^s, p^o but NOT p^w) can be stated in terms of n^s and n^o. Thus, the profit function has been manipulated in such a way that the problem is restated as a three-dimensional optimisation with respect to (n^s, n^o, p^w) instead of (r, n, p^w). The total derivative of platform's profits Π with respect to p^w can be written as

$$\begin{aligned}\frac{d\Pi}{dp^w} &= \frac{\partial\Pi}{\partial n^s}\frac{dn^s}{dp^w} + \frac{\partial\Pi}{\partial n^o}\frac{dn^o}{dp^w} + \frac{\partial\Pi}{\partial p^w} \\ &= \frac{\partial\Pi}{\partial n^s}\frac{dn^s}{dp^w} + \frac{\partial\Pi}{\partial n^o}\frac{dn^o}{dp^w} - D(\mathbf{p})(n^sq^o_{im} + \sum_{j=n^s+1}^{n^s+n^o} q^o_j)\end{aligned}$$

We also know that in the three-dimensional optimisation, the partial derivatives of $\Pi(n^s, n^o, p^w)$ with respect to n^s and n^o are both zero[19] (i.e. $\frac{\partial\Pi}{\partial n^s} = 0$ and $\frac{\partial\Pi}{\partial n^o} = 0$). Thus, we obtain

$$\frac{d\Pi}{dp^w} = -D(\mathbf{p})(n^sq^o_{im} + \sum_{j=n^s+1}^{n^s+n^o} q^o_j) < 0$$

The negative sign of the above derivative implies that the platform's profits, Π, decrease in p^w. Thus, it implies that the profit maximisation for the platform necessarily requires to have p^w at its lower bound, which is at the marginal cost c. ■

Proof of Corollary 2.3. Recall from the equivalence of the platform's profit function (2.11)

[19]The partial derivatives are zero when there is an interior solution.

and (2.23), we have

$$\begin{aligned}\Psi(n^s, n^o) &= D(n^s, n^o)(n^s q^s p^s + n^o q^o p^o - c) \\ &= D(n^s, n^o)[p^o(n^s, n^o) - \mu - c] \end{aligned} \quad (2.30)$$

Since $D(n^s, n^o)$ and $p^o(n^s, n^o)$ are both endogenous in n^s, these two variables are simultaneously affected by the change in n^s, i.e. from $E_{n^s+n^o}$ to $E'_{n^s+n^o+1}$. To scrutinise how the equilibrium value of expression (2.30) is affected, we first look at the simplified *asymmetric* pricing mechanism by fixing the existing prices of the $(n^s + n^o)$ brands. In another word, the only choice variable is the price of the $(n^o + 1)$th platform merchandised brand. Here we put a subscript f to denote the variables in this case. Then we have

$$\Psi_f(n^s, n^o + 1) = D_f(n^s, n^o + 1)[(n^s q_f^s p_f^s + n^o q_f^o p_f^o - c) + q^o_{f,n^o+1} p^o_{f,n^o+1}]$$

We also know from (2.13) and (2.14) that

$$\lim_{p^o_{f,n^o+1} \to \infty} q^o_{f,n^o+1} = 0$$

$$\lim_{p^o_{f,n^o+1} \to \infty} D_f(n^s, n^o + 1) = D_f(n^s, n^o)$$

Thus, we can conclude that $\Psi_f(n^s, n^o + 1) = \Psi(n^s, n^o)$ when p^o_{f,n^o+1} approaches infinity, which means that $\Psi_f(n^s, n^o+1)$ at $E'_{n^s+n^o+1}$ can be as high as $\Psi(n^s, n^o)$ at $E_{n^s+n^o}$ by making the $(n^o + 1)$th brand priceless. This result is intuitive in that the existing market profile will not be altered if the additional $(n^o + 1)$th brand charges a ridiculously high price.

It implies that by imposing an asymmetric pricing strategy for the platform, the value Ψ at $E'_{n^s+n^o+1}$ is at least as high as that at $E_{n^s+n^o}$. We also note that $p_f^o < p^o_{f,n^o+1} = \infty$ where $j = 1, ..., n^o$, whereas in equilibrium all the prices should be symmetric. Therefore, we can conclude that Ψ is **strictly** preferred at $E'_{n^s+n^o+1}$ to $E_{n^s+n^o}$.

Here we have completed the proof of Corollary 2.3. ■

Proof of Corollary 2.2 and Corollary 2.4. Firstly, we prove the second statement in Corollary 2.2 that Ψ is concave. Since it is difficult to directly show a negative second derivative to justify concavity of the Ψ function, we try to approach the answer by looking at the sign of $\frac{d\Psi(n^s, \bar{n}-n^s)}{dn^s}|_{n^s \to \bar{n}}$.

We obtain a more extensive form of the Ψ function by expanding the expression of $D(n^s, \bar{n} - n^s)$. That is

$$\Psi(n^s, \bar{n} - n^s) = [1 - e^{-(\bar{n}-n^s)e^{(v-p^o)/\mu} - n^s e^{(v-p^s)/\mu)}}](p^o - \mu - c).$$

We obtain the first derivative by differentiating the above expression with respect to n^s.

That is

$$
\begin{aligned}
& \frac{d\Psi(n^s, \bar{n}-n^s)}{dn^s} \\
= \; & -e^{-M}e^{v/\mu}(e^{-p^o/\mu} + \frac{\bar{n}-n^s}{\mu}e^{-p^o/\mu}\frac{dp^o}{dn^s} - e^{-p^s/\mu} + \frac{n^s}{\mu}e^{-p^s/\mu}\frac{dp^s}{dn^s}) \\
& \cdot(p^o - \mu - c) + (1-e^{-M})\frac{dp^o}{dn^s} \qquad (2.31) \\
= \; & -e^{-M}e^{v/\mu}(e^{-p^o/\mu} - e^{-p^s/\mu})(p^o - \mu - c) - e^{-M}e^{(v-p^s)/\mu}\frac{n^s}{\mu}\frac{dp^s}{dn^s} \\
& +[1 - e^{-M} - (p^o - \mu - c)e^{-M}e^{(v-p^o)/\mu}\frac{\bar{n}-n^s}{\mu}]\frac{dp^o}{dn^s} \qquad (2.32) \\
\text{where } M = \; & (\bar{n}-n^s)e^{(v-p^o)/\mu} + n^s e^{(v-p^s)/\mu}
\end{aligned}
$$

When $n^s \to \bar{n}$, (2.18) yields that $p^s \to \frac{\mu}{1-1/\bar{n}} + c$; and (2.21) yields that $p^o \to (\frac{\mu}{1-1/\bar{n}} + c + \mu)$. Furthermore, it implies from (2.22) that $\frac{d(p^o - p^s)}{dn^s} \to -\frac{e^{-1}\mu}{\bar{n}}$; while from (2.18) that $\frac{dp^s}{dn^s} \to -\frac{(1-e^{-1})\mu}{(\bar{n}-1)^2}$. Thus, we conclude that $\frac{dp^o}{dn^s} \to -[\frac{e^{-1}\mu}{\bar{n}} + \frac{(1-e^{-1})\mu}{(\bar{n}-1)^2}]$.

Then, the first derivative of Ψ (2.31) when $n^s \to \bar{n}$ becomes

$$
\begin{aligned}
\frac{d\Psi(n^s, \bar{n}-n^s)}{dn^s}|_{n^s \to \bar{n}} = \; & -\exp(v/\mu - \bar{n}e^{(v-p^s)/\mu})(e^{-p^o/\mu} - e^{-p^s/\mu} + \frac{\bar{n}}{\mu}e^{-p^s/\mu}\frac{dp^s}{dn^s})(p^o - \mu - c) \\
& +[1 - \exp(-\bar{n}e^{(v-p^s)/\mu})]\frac{dp^o}{dn^s}
\end{aligned}
$$

Substituting for p^o, p^s, $\frac{dp^o}{dn^s}$ and $\frac{dp^s}{dn^s}$ in the above expression, we can obtain a function of $\bar{n}$ with parameter μ, c and v as follows (We define that $\phi(\bar{n}) = \frac{d\Psi(n^s, \bar{n}-n^s)}{dn^s}|_{n^s \to \bar{n}}$).

$$
\begin{aligned}
\phi(\bar{n}) \equiv \; & \frac{d\Psi(n^s, \bar{n}-n^s)}{dn^s}|_{n^s \to \bar{n}} \\
= \; & -\exp(v/\mu - \bar{n}e^{-\frac{1}{1-1/\bar{n}} - \frac{c}{\mu} + \frac{v}{\mu}})[\exp(-\frac{1}{1-1/\bar{n}} - 1 - \frac{c}{\mu}) - \exp(-\frac{1}{1-1/\bar{n}} - 1) \\
& -\frac{\bar{n}}{\mu}\exp(-\frac{1}{1-1/\bar{n}} - 1)\frac{(1-e^{-1})\mu}{(\bar{n}-1)^2}]\frac{\mu}{1-1/\bar{n}} \\
& -[1 - \exp(v/\mu - \bar{n}e^{-\frac{1}{1-1/\bar{n}} - \frac{c}{\mu} + \frac{v}{\mu}})](\frac{e^{-1}\mu}{\bar{n}} + \frac{(1-e^{-1})\mu}{(\bar{n}-1)^2}) \qquad (2.33)
\end{aligned}
$$

Recall that $\bar{n} \in [2, \infty)$, otherwise, the reaction functions will be undefined for the market competitors. We can show that $\phi(\bar{n}) = \frac{d\Psi(n^s, \bar{n}-n^s)}{dn^s}|_{n^s \to \bar{n}} > 0$ and is finite when $\bar{n} = 2$; and $\frac{d\phi(\bar{n})}{d\bar{n}} < 0$ when $\bar{n} \geq 2$.(See the illustration in Figure 2.10, and the detailed proof is to be provided). Now we can conclude that the Ψ function is concave.

Secondly, by using $\phi(\bar{n}) = \frac{d\Psi(n^s, \bar{n}-n^s)}{dn^s}|_{n^s \to \bar{n}}$ and $\frac{d\phi(\bar{n})}{d\bar{n}} < 0$ as derived in the previous paragraphs, now it is straightforward to yield Corollary 2.4.

Finally, we need to go back to prove the first statement in Corollary 2.2 to show the slope of Ψ function at $n^s = 1$ is always positive.

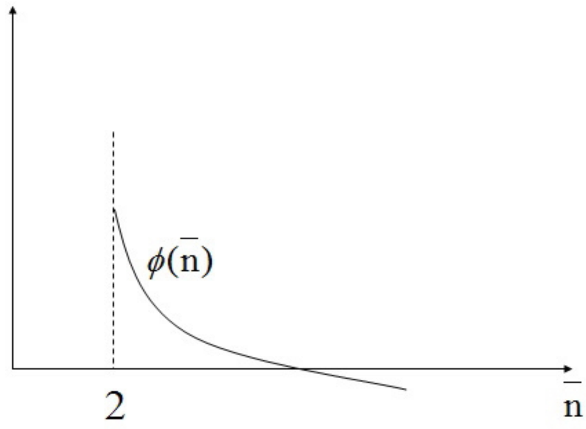

Figure 2.10: Illustration of $\phi(\bar{n})$ function

By using (2.32) the slope of the Ψ function at $n^s = 1$ can be obtained as follows.

$$\begin{aligned} & \frac{d\Psi(n^s, \bar{n} - n^s)}{dn^s}|_{n^s=1} \\ = & -e^{-M} e^{v/\mu}(e^{-p^o/\mu} - e^{-p^s/\mu})(p^o - \mu - c) - e^{-M} e^{(v-p^s)/\mu} \frac{1}{\mu} \frac{dp^s}{dn^s} \\ & +[1 - e^{-M} - (p^o - \mu - c)e^{-M} e^{(v-p^o)/\mu} \frac{\bar{n} - 1}{\mu}] \frac{dp^o}{dn^s} \qquad (2.34) \\ \text{where } M = & (\bar{n} - 1)e^{(v-p^o)/\mu} + e^{(v-p^s)/\mu} \end{aligned}$$

We learn from Corollary 2.4 that at any $n^s \in [1, \bar{n}]$ the slope of the Ψ function becomes smaller as $\bar{n}$ increases. Thus, we can prove the second statement in Corollary 2.2 by showing a positive slope of the Ψ function at $n^s = 1$ when $\bar{n} \to \infty$. The proof proceeds in this way.

$p^o > p^s$ yields that $(e^{-p^o/\mu} - e^{-p^s/\mu}) < 0$, thus the first term in (2.34) is positive.

We also learn from Lemma 2.2 (i) that $\frac{dp^s}{dn^s} < 0$, thus the second term in (2.34) is also positive.

Since $\frac{dp^o}{dn^s}$ in the last term in (2.34) is negative (Lemma 2.2 (i)), the proof is complete if we can show the following inequality as $\bar{n} \to \infty$.

$$1 - e^{-M} - (p^o - \mu - c)e^{-M} e^{(v-p^s)/\mu} \frac{\bar{n} - 1}{\mu} < 0$$

$$\Leftrightarrow e^M < 1 + (p^o - \mu - c)\frac{M - e^{(v-p^s)/\mu}}{\mu} \text{ where } M = (\bar{n} - 1)e^{(v-p^o)/\mu} + e^{(v-p^s)/\mu}$$

It can be implied from (2.27) and (2.28) that $p^o \to \infty$ and $M \to \infty$ when $\bar{n} \to \infty$. Thus the above inequality is true. Here we have completed all the proofs of Corollary 2.2 and Corollary 2.4. ■

Proof of Lemma 2.3. Let us recall the derivative of the Ψ function as (2.32), which can be rewritten as

$$\begin{aligned}\frac{d\Psi(n^s,\bar{n}-n^s)}{dn^s} &= -e^{-M}e^{v(\bar{n})/\mu}(e^{-p^o/\mu}-e^{-p^s/\mu})(p^o-\mu-c)-e^{-M}e^{(v(\bar{n})-p^s)/\mu}\frac{n^s}{\mu}\frac{dp^s}{dn^s} \\ &\quad +(1-e^{-M})\frac{dp^o}{dn^s}-(p^o-\mu-c)e^{-M}e^{(v(\bar{n})-p^o)/\mu}\frac{\bar{n}-n^s}{\mu}\frac{dp^o}{dn^s} \qquad (2.35) \\ \text{where } M &= (\bar{n}-n^s)e^{(v(\bar{n})-p^o)/\mu}+n^s e^{(v(\bar{n})-p^s)/\mu}\end{aligned}$$

Firstly, we note that the implicit reaction functions such as (2.27) and (2.28) are irrelevant to $v(\bar{n})$, thus all the elements in (2.35) with regard to p^o and p^s are unchanged if indirect network effect is brought in. Secondly, we need to note that $v(\bar{n})$ affects both M and $e^{v(\bar{n})/\mu}$ in the above expression. Recall that $(e^{-p^o/\mu}-e^{-p^s/\mu})$, $\frac{dp^s}{dn^s}$ and $\frac{dp^o}{dn^s}$ are all negative, thus a sufficient condition to yield $\frac{\partial^2}{\partial n^s \partial v(\bar{n})}\Psi(n^s,\bar{n}-n^s)<0$ is

$$\begin{aligned}\frac{d(e^{-M}e^{v(\bar{n})/\mu})}{dv(\bar{n})} &= \frac{d(e^{-M}e^{v(\bar{n})/\mu})}{de^{v(\bar{n})/\mu}}\frac{de^{v(\bar{n})/\mu}}{dv(\bar{n})}<0 \\ &\Leftrightarrow \frac{d(e^{-M}e^{v(\bar{n})/\mu})}{de^{v(\bar{n})/\mu}}<0 \Leftrightarrow M>1\end{aligned}$$

Thus, when there is indirect network effect in consume demand and $M > 1$, the slope of the "gross profits" curve Ψ becomes smaller and the optimal portfolio tends to have more platform merchandised brands (i.e. decrease in n^s) at every level of brand variety. Here we have completed the proof of Lemma 2.3. ■

Chapter 3

Hotelling Competition in Marriage Markets

Abstract

Motivated by the Chinese experience, this chapter develops a two-sided matching model of marriage markets with search frictions. In particular, we extend the one-market model of the previous literature to a two-market one, allowing the agents to migrate between the markets at a fixed cost. The analysis focuses on the unmatched and the migrating population which is induced by the different initial sex ratios in different marriage markets. We characterise the equilibrium set in the parameter space of migrating costs. Our results suggest that the large scale of male migration from rural to urban areas can be explained by its relatively lower migrating costs and the higher sex-ratio-at-birth (boys to girls) in rural areas. We also explore the welfare implications and find that a full subsidy of the migrating costs does not necessarily benefit those who receive them but always benefits the opposite sex if they are the short sides of both markets.

3.1 Introduction

We model two-sided matching markets with unbalanced inflows and homogeneous agents. One idea about matching with unbalanced entry is that the vertical heterogeneity in quality results in an assortative matching format in equilibrium (Chade, 2001). It means that all the active agents in the market end up with positive utility, leaving the rest of the market empty-handed[1]. In this chapter we set up an alternative market (which means we have two linked markets in the model) where agents can try to obtain higher (expected) utility. This is typically the case for the marriage market, job market or even trade in real life, where a proportion of the agents

[1] This proportion of agents will simply choose not to enter the market in the first place if we assume the utility of remaining unmatched is zero.

will choose to look for potential partners in an alternative market in which they are more likely to get a higher utility, instead of remaining unmatched in their own market. While it is not straightforward to see how the agents with high quality move in equilibrium, it is relatively easy to comprehend the bottom agents' motivation for looking for partners in an alternative market, rather than otherwise being left out in his/her birthplace market. However, the issue is complicated with heterogeneity in the two-market model since different patterns of externalities would be brought about by the inter-flows, and at the same time, it would alter the distribution of the stocks. Therefore, we will only discuss the case with homogeneous agents in this chapter.

The setup of the unbalanced inflows is motivated by the fact that, in many East or South Asian countries, the sex ratios *at birth* (i.e. boys to girls) are, more often than not, greater than unity. Oster (2005) quotes that the *population* sex ratio is 1.07 in China, 1.08 in India, and 1.11 in Pakistan. Due to the relatively higher mortality of the new-born boys, as well as the lower average life expectancy of men, the sex ratios at birth in these countries turn out to be even higher. Moreover, since the technology of type-B ultrasound (which helps to identify the sex of the fetuses) became available in the mid-1980s, the unbalanced sex ratios at birth become more severe due to induced abortion. For China, the culture of son preference, as well as the differentiated family planning policies for different regions and ethnic groups to control the population[2] mean that there are several characteristics in the distribution of the unbalanced sex ratios *at birth*. For example, according to the Fifth National Population Census of China (2002), the sex-ratio-at-birth in rural areas of China is 1.21, whilst in urban areas this is 1.14.[3]

With the unbalanced numbers of entrants into the marriage markets in our model, we allow the singles to seek partners not only in his/her birthplace market but also in the other market at some additional migrating cost. We are interested in the analysis of the migration between markets, as well as the unmatched populations in each market in equilibrium. We investigate the following questions: Who moves in equilibrium? What is the correlation between the migrating population in equilibrium and the migrating costs? Will welfare be enhanced if the government subsidises the migrating population to offset the moving costs? If yes, who benefits from the policy of subsidies? As a matter of fact, data shows that since the mid-1980's China has witnessed the largest rural to urban labour flow in the world's history, of which the majority are male. Notably, a commonly cited figure puts the number of rural migrants residing in urban areas at 50 million in the mid-1990's.

The rural-to-urban migration can be attributed to the urbanisation drive as well as the relaxation of long-standing policies against rural-to-urban migration, or more specifically, the incentive of "getting a better job" (Lall, et al, 2006) as discussed in other literature. However,

[2] For instance, the "One Child Policy" applies only to the majority Han Chinese in urban areas but not those in rural areas, nor the minority ethnic groups.

[3] Oster (2005) established and later retracted her hypothesis that women with hepatitis B are more likely to give birth to male children than female ones.

there are other incentives that can explain the migration, for example the motivation to search for a better mate by maximising his/her expected lifetime utility from marriage.

The arguments are developed in the context of the two-sided matching model of a marriage market, whose settings typically feature heterogeneous agents and costly search, e.g. the discounting in Burdett and Coles (1997), and the fixed search cost in Chade (2001) and Atakan (2006). The random searching in the heterogeneous case reflects the fact that one individual cannot tell the quality of others before they meet. However, since in the model we focus mainly on migration between markets in the model, we assume all the agents are homogeneous in order to minimise complications. This also implies that the matching set of any single (the set of agents with whom that single is willing to match) is simply the whole population of the opposite side market. In other words, any two agents from different sides of the market are mutually acceptable, and they would be immediately matched upon meeting. Thus there are no search issues, and one simply meets the other agents randomly.

There is numerous related literature on the one-market matching model of marriage. The most frequently referred literature, Becker's papers (1973,1974), show that in a matching market with transferable utility (which means the successfully matched agents can fully divide their joint output) and supermodularity (which means that there is complementarity in joint production), every competitive equilibrium exhibits positive assortative matching (PAM). That is to say that the best man marries the best woman, the second-best man marries the second-best woman, and so on. Furthermore, PAM also automatically ensures the maximisation of output; and therefore welfare if we calibrate welfare in terms of output.

However in Becker's model there is no friction involved in searching for a desirable matching mate. Subsequent researchers introduce search friction - such as time cost and fixed search cost - and also take the rate of random meeting into consideration. Based on these settings, Sattinger (1995), Lu and McAfee (1996) and Chade (2001) consider transferable utility (where the matched agents split the output by Nash Bargaining) with ex ante heterogeneity, as well as endogenising the unmatched distribution. Shimer and Smith (2000) prove that within the framework of a discounted search model, complementarities in joint production are not sufficient for positive assortative matching. The intuition here is that the agents need to pay for the search friction, which lowers the probability of being matched with a desirable partner without friction. Consequently, the high quality agent might end up settling with a low quality agent, and thus the assortative matching breaks down. Atakan's model (2006) includes transferable utility and constant additive costs, showing that assortative matching relies on the Constant Surplus Condition[4], which asserts that every agent enjoys the same expected surplus in the future match.

[4]This condition emerges since the expected surplus from future matches is the benefit of additional search and must, at optimum, equal to the constant cost of search.

Under the conditions of nontransferable utility, search friction and heterogeneity, Burdett and Coles (1997) and Smith (2006) assume that the utility derived from marriage increases in his/her partner's quality. They show that the equilibrium is characterised by a class partition (or "perfect segmentation" called by Smith), which is to say that marriages are formed only between couples that belong to the same class/segment. Chade (2001) presents a model with non-transferable utility and constant additive costs.

Burdett and Coles's model (1997) of a two-sided matching market with quality heterogeneity (without migration costs) is a direct forbearer of our analysis. Under search friction (i.e. time cost by discounting) and the non-transferable utility assumption in our model, the results of their paper are similar to ours in the one market context. We extend the much discussed one-market searching model to a two-market one at the expense of losing generality of heterogeneity to homogeneity. In our model, the new entrants are provided an alternative market to search for partners, but migrating from the birthplace to the other requires a fixed cost. In the process of matching, the agents search over an infinite horizon for a partner to maximise their discounted payoffs, which is different from the undiscounted payoff assumption in Atakan's model. We also adopt the non-transferable utility assumption and assume that each agent's utility derived from the marriage equals the quality of her/his partner.

In this chapter, we mainly analyse the migrating population and the unmatched population in equilibrium when the outflows are exactly balanced by the inflows of new entrants. Our result shows that in the settings of the two-market model with homogeneous agents, discount rate and fixed migrating costs, there are four types of equilibrium (i.e. either men or men migrate, both men and women migrate, and neither men or women migrate) in the parameter space of migrating costs. Moreover, we explore the welfare implication within the framework of this two-market model. The issue of the aggregate welfare turns out to be too complicated to solve analytically due to the different patterns of externalities brought about by migration. However, we find that the aggregate welfare of a single side as a function of its gender inflow ratio is closely related to the concavity of the expected utility of joining a market. Our result also shows that the full subsidies of migrating costs do not necessarily benefit those who receive them, but do always benefit the opposite sex who are the short sides of the markets.

This chapter proceeds in following way: Section 3.2 outlines the general model. Section 3.3 discusses the general description of equilibria in the parameter space of the migrating costs and also includes some numerical examples. Section 3.4 provides a discussion of the welfare implications. Section 3.5 concludes.

3.2 The Model

Market Setup: It is a continuous time two-sided matching model with search friction, and there are a large number of unmatched men and women searching potential mates in the marriage markets. The reason for such friction lies in the fact that meeting/dating with other agents is time-consuming and meeting Mr/Miss Right is haphazard. Each agent randomly meets the rest of the market, which intuitively makes the unmatched population much larger than the case without random meeting. The reason why there are unmatched individuals is either that they simply have not met yet or they have met but are not mutually acceptable. The former case is determined by meeting technology, and the latter by matching set. For the sake of simplicity, we assume that all agents have unit qualities.

We assume that there are two geographically isolated marriage markets, e.g. the urban and rural marriage markets, which are denoted by u and r for short. At any given moment, there is exogenous initial inflow created in each marriage market[5]. The population of the exogenous instantaneous inflow of sex j in market p is denoted by I_j^p, where $p = u, r$ and $j = m, w$. Accordingly, the unmatched population of sex j in the marriage market p is denoted by S_j^p; the *migrating* population from urban to rural areas by ΔI_j^{ur}, and from rural to urban areas by ΔI_j^{ru}.

Utility Function: We assume that marriage gives utility 1, and being single has utility 0.

Strategies: In Burdett and Coles (1997), agents immediately observe each other's quality upon meeting. If both propose to each other, they form a marriage and leave the market; if at least one vetoes, they ignore each other and continue to look for other partners. In our model of homogeneous agents, all the agents from different sides of the market are mutually acceptable and they will form a marriage upon meeting. For the individuals who comprise the initial inflows, they always have a *one-shot* choice of either entering their birthplace market or the alternative market at a fixed cost of migrating. Here, the migrating cost for men (women) from rural to urban area is denoted by C_m^{ru} (C_w^{ru}), and from urban to rural area by C_m^{ur} (C_w^{ru}). Each agent maximises his/her expected present value of payoffs, discounted at the rate ρ. In the procedure of deciding which market to enter, the newly matured agent evaluates his/her expected present value in both marketplaces. Once the choice has been made, they form the *ultimate* inflows which are indicated by putting primes to the *initial* inflows, i.e. $I_j'^p$. Together with those already in the market, the agents will be restricted to seeking desirable partners in this chosen marketplace. We also assume that the successful matches leave the market and never split nor return to the marriage market again. Moreover, in the setup of continuous time model, we assume that some agents leave the market due to physical reasons, e.g. death or being too old to get married, and we denote the instantaneous probability of non-marital leaving by δ.

Meeting Technology: Each agent randomly meets the rest of the market in pairs. Here, we use constant-return-to-scale meeting technology. Time is continuous, and at each moment an

[5] The initial inflows can be interpreted as the newly matured agents entering the marriage markets; and they might be further set as endogenous in more general cases.

agent of sex j meets a single agent of the opposite sex with probability $\alpha_j(S^p_{-j}, S^p_j)$. In order to equate the measure of men who meet women and that of women who meet men, we assign an explicit form of the meeting probability:

$$\alpha_j(S^p_{-j}, S^p_j) = \alpha_0 S^p_{-j}/(S^p_j + S^p_{-j})$$

where $j = m, w$, and $\alpha_0 \in (0, 1)$ is the meeting efficiency parameter.

Equilibrium: In stationary equilibrium, the inflow must exactly replicate the outflow in quantity (and quality) on each side of each market. As described earlier, the outflows are comprised of two parts: the agents who are successfully matched at each moment and then leave the market permanently, which is $S^p_j \cdot \alpha_j(S^p_{-j}, S^p_j) = \alpha_0 S^p_j S^p_{-j}/(S^p_j + S^p_{-j})$, plus the agents who are forced out of the market by physical reasons with instantaneous probability $\delta > 0$. Putting them together, in equilibrium we have

$$I'^p_j = \frac{\alpha_0 S^p_j S^p_{-j}}{S^p_j + S^p_{-j}} + S^p_j \cdot (1 - \frac{\alpha_0 S^p_{-j}}{S^p_j + S^p_{-j}})\delta.$$

3.3 Equilibrium Analysis

We start by characterising the stationary equilibrium and will investigate the condition for its existence later.

Let W^p_w denote the *expected* present value of an unmatched woman who is still searching a potential mate in market p. While unmatched, she earns nothing, but she meets single men at an instantaneous probability $\alpha_0 S^p_m/(S^p_m + S^p_w)$ if she seeks partners in marriage market p. Here, we assume that unattached agents perfectly estimate the meeting probability with the opposite sex by observing the unmatched sex ratio in the market. As a result, each agent employs a stationary threshold strategy in equilibrium. The dynamic programming equation implies that

$$\begin{aligned}
\frac{1+\rho}{1-\delta} W^p_w &= \frac{\alpha_0 S^p_m}{S^p_m + S^p_w} + (1 - \frac{\alpha_0 S^p_m}{S^p_m + S^p_w}) W^p_w \\
&\Rightarrow \; W^p_w = \frac{\alpha_0 \theta S^p_m}{S^p_w + S^p_m + \alpha_0 \theta S^p_m} \qquad (3.1) \\
\text{similarly } W^p_m &= \frac{\alpha_0 \theta S^p_w}{S^p_m + S^p_w + \alpha_0 \theta S^p_w}, \qquad (3.2) \\
\text{where } \theta &= \frac{1-\delta}{\rho + \delta}.
\end{aligned}$$

Before we turn to the two-market model, let us first introduce a proposition concerning a single homogeneous market.

Proposition 3.1 *Given meeting efficiency α_0, positive market leaving rate δ,*[6] *inflows I_m and I_w in a single homogeneous market, there always exists a unique stationary equilibrium (SE). That is*

$$\begin{aligned} S_w &= \frac{2\delta I_w - (I_m - I_w)(\alpha_0 - \alpha_0\delta + \delta) + \sqrt{D}}{2\delta(\alpha_0 - \alpha_0\delta + 2\delta)} \\ S_m &= \frac{2\delta I_m - (I_w - I_m)(\alpha_0 - \alpha_0\delta + \delta) + \sqrt{D}}{2\delta(\alpha_0 - \alpha_0\delta + 2\delta)}, \\ \textit{where } D &= (I_m - I_w)^2(\alpha_0 - \alpha_0\delta + \delta)^2 + 4I_m I_w \delta^2. \end{aligned}$$

Proof. See Appendix. ■

Proposition 3.1 shows that the unmatched population of each sex in each market (S_j) is uniquely determined by the inflows (I_j). In the following paragraphs, we extend the one-market model in Proposition 3.1 to fit into a two-market model with homogeneous agents, given meeting efficiency α_0, positive market leaving rate δ, and *initial* inflows I_m^u, I_w^u, I_m^r, I_w^r. Without loss of generality, we assume throughout the discussion that the male-to-female sex ratio is higher in the rural area than that in the urban area[7], i.e. $I_m^u/I_w^u < I_m^r/I_w^r$; the relevant migrating costs $C_w^{ru} = C_m^{ur} = \infty$, C_w^{ur} and C_m^{ru} are positive. Next, we introduce Proposition 3.2 which shows some basic features concerning the migrating population between the two homogeneous markets.

Proposition 3.2 *(i) For the same side of the two markets, the migration is unilateral rather than bilateral;*

(ii) if both sides of the markets move, the migrating directions are opposite.

Proof. The proof of (i) is straightforward. The marginal benefit of entering an alternative market only derives from a higher rate of meeting the opposite sex which is endogenously determined by the *ultimate* inflows into each market. Recall that the preferences of all agents are identical, therefore if it is beneficial for an agent to move from his/her birthplace market to another, it can never be beneficial for any of his/her competitors in the other market to move in the opposite direction.

Next, we prove (ii) by contradiction. Without loss of generality, we suppose that there are both urban-born men and women moving to rural marriage market in SE. Thus, in equilibrium the marginal benefit for the urban-born agent to move must be equal to the marginal migrating

[6] It is shown in the Appendix that *positive* market leaving rate is a necessary condition for the equilibrium with unbalanced inflows.

[7] As we have described in the introduction, it is especially the case in China, e.g. the sex-ratio-at-birth is 1.21 in rural areas and 1.14 in urban areas.

cost incurred. That is

$$\begin{aligned}\frac{\alpha_0\theta S^r_w}{S^r_w+S^r_m+\alpha_0\theta S^r_w}-\frac{\alpha_0\theta S^u_w}{S^u_w+S^u_m+\alpha_0\theta S^u_w} &= C^{ur}_m=\infty \\ \text{(for urban-born men)} &\implies \frac{S^r_w}{S^r_m}\gg\frac{S^u_w}{S^u_m}\end{aligned} \tag{3.3}$$

$$\begin{aligned}\frac{\alpha_0\theta S^r_m}{S^r_w+S^r_m+\alpha_0\theta S^r_m}-\frac{\alpha_0\theta S^u_m}{S^u_w+S^u_m+\alpha_0\theta S^u_m} &= C^{w}_{ur}\geq 0 \\ \text{(for urban-born women)} &\implies \frac{S^r_m}{S^r_w}\geq\frac{S^u_m}{S^u_w}\end{aligned} \tag{3.4}$$

We note that inequalities (3.3) and (3.4) contradict each other. Therefore, in SE there must not exist two groups of agents simultaneously migrating in the same direction. Here we have completed the proofs of Proposition 3.2. ■

As a matter of fact, the intuition of Proposition 3.2 is simply that whenever one side of the markets is on the run for higher expected payoffs, the marginal benefit to the agents on the other side of moving in the same direction must be negative.

In order to analyse the migrating population, or the incentive for men and/or women to migrate, we need to answer how the sex ratios in the birthplace and alternative market are related to their expected payoffs. Next, we establish a corollary and a lemma to answer this question. Here, the ultimate inflows are defined as the inflows when the men and women have made up their minds of entering the target markets.

Corollary 3.1 *Given the meeting efficiency α_0, the market leaving rate δ and the initial inflow I^p_j, a unique equilibrium is guaranteed when the sex ratios of the **ultimate** inflows, i.e. $\Gamma^p_j = I'^p_{-j}/I'^p_j$, are known.*

Proof. Based on Proposition 3.2, we suppose there are two opposite inter-flows between the markets, i.e. ΔI^{ru}_m and ΔI^{ur}_w. By definition, we have

$$\frac{I^u_m+\Delta I^{ru}_m}{I^u_w-\Delta I^{ur}_w}=\frac{I'^u_m}{I'^u_w}=\Gamma^u_w$$

$$\frac{I^r_m-\Delta I^{ru}_m}{I^r_w+\Delta I^{ur}_w}=\frac{I'^r_m}{I'^r_w}=\Gamma^r_w$$

After a bit more manipulation, we obtain

$$\Delta I^{ru}_m=\frac{-\Gamma^r_w I^u_m+\Gamma^u_w\Gamma^r_w I^u_w-\Gamma^u_w I^r_m+\Gamma^u_w\Gamma^r_w I^r_w}{\Gamma^r_w-\Gamma^u_w}$$

$$\Delta I^{ur}_w=\frac{I^u_m-\Gamma^u_w I^u_w+I^r_m-\Gamma^r_w I^r_w}{\Gamma^r_w-\Gamma^u_w}$$

We note from the above two equations that ΔI^{ru}_m and ΔI^{ur}_w are uniquely determined by Γ^u_w

and Γ^r_w, thus, the four *ultimate* inflows I'^p_j can be expressed in terms of Γ^u_w and Γ^r_w. Recalling the statement in Proposition 3.1 that a unique equilibrium is guaranteed by the *ultimate* (or the realised) inflows into the markets, we can now conclude that the unique equilibrium is also guaranteed by knowing Γ^u_w and Γ^r_w. Therefore, we complete the proof of Corollary 3.1. ■

Corollary 3.1 indicates that the unmatched population S^p_j can be estimated either by observing the sizes of migrating population, i.e. ΔI^{ru}_m and ΔI^{ur}_w, or by knowing the sex ratios of the *ultimate* inflows, i.e. Γ^u_w and Γ^r_w.

In the following paragraphs, we will discuss the correlation among W^p_j (the expected present value for the sex j agents in market p), M^p_j (which is defined as the sex ratio of unmatched *population* in marriage market p, i.e. $M^p_j = S^p_{-j}/S^p_j$), and $\Gamma^p_j = I'^p_{-j}/I'^p_j$ (the sex ratio of the *ultimate* inflows as defined in Corollary 3.1) in stationary equilibrium.

Lemma 3.1 *(i)* M^p_j *is a strictly increasing function of* Γ^p_j*;*

(ii) $M^p_j \gtreqless \Gamma^p_j$ *if and only if* $\Gamma^p_j \gtreqless 1$*; and* $M^p_j = \Gamma^p_j = 1$ *if and only if* $\Gamma^p_j = 1$*;*

(iii) W^p_j *is a strictly increasing function of* Γ^p_j*, i.e.* $W^p_j = f(\Gamma^p_j)$*, and bounded above by* $\dfrac{\alpha_0\theta}{1+\alpha_0\theta}$.

Proof. First, let us look at the first statement. Since we have assumed $I^u_m/I^u_w < I^r_m/I^r_w$, the only scenario that can arise in equilibrium is that men migrate from rural to urban area and/or women from urban to rural area, if any; i.e. $\Delta I_m = \Delta I^{ru}_m = -\Delta I^{ur}_m$, $\Delta I_w = \Delta I^{ur}_w = -\Delta I^{ru}_w$, where ΔI_m and ΔI_w are both non-negative. As the outflows replicate the ultimate inflows in equilibrium, the following conditions must be satisfied:

$$(1-\delta)\frac{\alpha_0 S^u_m S^u_w}{S^u_w + S^u_m} + \delta S^u_m = I'^u_m = I^u_m + \Delta I_m \tag{3.5}$$

$$(1-\delta)\frac{\alpha_0 S^u_m S^u_w}{S^u_w + S^u_m} + \delta S^u_w = I'^u_w = I^u_w - \Delta I_w \tag{3.6}$$

$$(1-\delta)\frac{\alpha_0 S^r_m S^r_w}{S^r_w + S^r_m} + \delta S^r_m = I'^r_m = I^r_m - \Delta I_m \tag{3.7}$$

$$(1-\delta)\frac{\alpha_0 S^r_m S^r_w}{S^r_w + S^r_m} + \delta S^r_w = I'^r_w = I^r_w + \Delta I_w \tag{3.8}$$

Dividing (3.6) by (3.5) and (3.8) by (3.7), we obtain

$$\frac{\alpha_0(1-\delta)M^p_j/(1+M^p_j) + M^p_j\delta}{\alpha_0(1-\delta)M^p_j/(1+M^p_j) + \delta} = \frac{I'^p_{-j}}{I'^p_j} = \Gamma^p_j$$

Manipulating the above equation, it yields that

$$\Gamma^p_j - 1 = \frac{M^p_j\delta - \delta}{\alpha_0(1-\delta)M^p_j/(1+M^p_j) + \delta}$$

$$\Rightarrow \frac{\alpha_0(1-\delta)(\Gamma_j^p - 1)}{\delta} = \frac{M_j^p - 1}{M_j^p/(1+M_j^p) + \delta/(\alpha_0 - \alpha_0\delta)} \tag{3.9}$$

We differentiate both sides of the above equation by Γ_j^p, and obtain

$$\begin{aligned} & \frac{\alpha_0(1-\delta)}{\delta} \\ = \quad & \frac{dM_j^p/d\Gamma_j^p[M_j^p/(M_j^p+1) + \delta/(\alpha_0 - \alpha_0\delta)] - (M_j^p - 1)/(M_j^p+1)^2 dM_j^p/d\Gamma_j^p}{[M_j^p/(M_j^p+1) + \delta/(\alpha_0 - \alpha_0\delta)]^2} \\ = \quad & \frac{dM_j^p/d\Gamma_j^p[(M_j^{p2}+1)/(M_j^p+1)^2 + \delta/(\alpha_0 - \alpha_0\delta)]}{[M_j^p/(M_j^p+1) + \delta/(\alpha_0 - \alpha_0\delta)]^2} > 0 \end{aligned}$$

$$\Rightarrow dM_j^p/d\Gamma_j^p > 0$$

Now we have shown that M_j^p is a strictly increasing function of Γ_j^p.

Next, let us turn to [ii] in the lemma.

We rewrite (3.9) in the following form.

$$M_j^p - 1 = (\Gamma_j^p - 1)[1 + \frac{\alpha_0(1-\delta)M_j^p}{\delta(1+M_j^p)}]$$

Since the last term in the square brackets is strictly greater than unity, it is now easy to arrive at the conclusion in Lemma 3.1 [ii].

Then, we move on to the last statement in the lemma.

From (3.1) and (3.2), the expected present value for the sex j agent in market p can be written as

$$\begin{aligned} W_j^p \quad &= \quad \frac{\alpha_0\theta S_{-j}^p}{S_j^p + S_{-j}^p + \alpha_0\theta S_{-j}^p} \\ &= \quad \frac{\alpha_0\theta}{S_j^p/S_{-j}^p + 1 + \alpha_0\theta} = \frac{\alpha_0\theta}{1/M_j^p + 1 + \alpha_0\theta} \end{aligned} \tag{3.10}$$

It can be easily implied from the above equation that W_j^p strictly increases in M_j^p. As we have already proved that M_j^p increases in Γ_j^p, we can now arrive at a well-defined function such that $W_j^p = f(\Gamma_j^p)$, which is also strictly increasing in Γ_j^p and bounded above by $\frac{\alpha_0\theta}{1+\alpha_0\theta}$.

Here we have completed all the proofs in Lemma 3.1 ■

Lemma 3.1 captures the intuition that the expected present value of a male (female) agent is positively related to the female-to-male sex ratio (the male-to-female sex ratio) of the *ultimate* inflows in the candidate market. More specifically, the sex ratio of the *ultimate* inflows determines that of the unmatched population, through which the expected present values of all agents are formed; in return, the decisions on entering which market affect the sex ratios of the *ultimate* inflows.

The result of Lemma 3.1 [ii] is also intuitive in that the imperfection of the matching technology necessarily leads to an increased sex ratio gap in the unmatched population, compared to that of the inflows.

In order to shed more light on the market profiles in equilibrium, we will establish Corollary 3.2 to Corollary 3.4 to help characterise the equilibrium set.

Let us define

$$\bar{C}_m^{ru} = f(I_w^u/I_m^u) - f(I_w^r/I_m^r) \quad (3.11)$$

$$\bar{C}_w^{ur} = f(I_m^r/I_w^r) - f(I_m^u/I_w^u) \quad (3.12)$$

We need to be careful that, unlike Lemma 3.1, $\bar{C}_m^{ru}$ and $\bar{C}_w^{ur}$ in the above definitions are determined by the sex ratio of the *initial* inflows, which are exogenous.

Corollary 3.2 *$C_m^{ru} \geq \bar{C}_m^{ru}$ and $C_w^{ur} \geq \bar{C}_w^{ur}$ are necessary conditions for the equilibrium in which neither side of the markets moves.*

Proof. If neither side of the market moves in equilibrium (that is, there are no inter-flows between markets), the necessary conditions are

$$W_m^u - W_m^r = f(\Gamma_m^u) - f(\Gamma_m^r) \leq C_m^{ru} \quad (3.13)$$

$$W_w^r - W_w^u = f(\Gamma_w^r) - f(\Gamma_w^u) \leq C_w^{ur} \quad (3.14)$$

$$\Gamma_m^u = I_w'^u/I_m'^u = I_w^u/I_m^u \quad (3.15)$$

$$\Gamma_m^r = I_w'^r/I_m'^r = I_w^r/I_m^r \quad (3.16)$$

Substituting for Γ_j^p in (3.13) and (3.14) by using (3.15) and (3.16), it follows that

$$f(I_w^u/I_m^u) - f(I_w^r/I_m^r) \leq C_m^{ru}$$

$$f(I_m^r/I_w^r) - f(I_m^u/I_w^u) \leq C_w^{ur}$$

Recalling the definitions of (3.11) and (3.12), we can conclude that for all $C_m^{ru} \geq \bar{C}_m^{ru}$ and $C_w^{ur} \geq \bar{C}_w^{ur}$ neither side of markets moves in equilibrium. Additionally, since the ultimate inflows are the same as the initial inflows in this case, we are able to derive the unmatched population S_j^p in equilibrium by using Proposition 3.1. ■

So far it is only shown that $C_m^{ru} \geq \bar{C}_m^{ru}$ and $C_w^{ur} \geq \bar{C}_w^{ur}$ are necessary for the non-migration equilibrium, and we will show that they are also sufficient conditions in the summary immediately after Corollary 3.4.

Now consider the situation that only men migrate from rural to urban area while women are *indifferent* between moving or not. Then, the following conditions must be satisfied in equilibrium.

$$W_m^u - W_m^r = f(\Gamma_m^u) - f(\Gamma_m^r) = f(1/\Gamma_w^u) - f(1/\Gamma_w^r) = C_m^{ru} \quad (3.17)$$

$$W_w^r - W_w^u = f(\Gamma_w^r) - f(\Gamma_w^u) = C_w^{ur} \tag{3.18}$$

$$\Delta I_m = \frac{-\Gamma_w^r I_m^u + \Gamma_w^u \Gamma_w^r I_w^u - \Gamma_w^u I_m^r + \Gamma_w^u \Gamma_w^r I_w^r}{\Gamma_w^r - \Gamma_w^u} > 0 \tag{3.19}$$

$$\Delta I_w = \frac{I_m^u - \Gamma_w^u I_w^u + I_m^r - \Gamma_w^r I_w^r}{\Gamma_w^r - \Gamma_w^u} = 0 \tag{3.20}$$

Intuitively, we are able to derive a function, i.e. $C_w^{ur} = h_1(C_m^{ru})$, from the equation system (3.17), (3.18) and (3.20). To see this, we substitute for Γ_w^r in (3.17) and (3.18) by using (3.20). That is

$$f(1/\Gamma_w^u) - f(\frac{I_w^r}{I_m^u - \Gamma_w^u I_w^u + I_m^r}) = C_m^{ru} \tag{3.21}$$

$$f(\frac{I_m^u - \Gamma_w^u I_w^u + I_m^r}{I_w^r}) - f(\Gamma_w^u) = C_w^{ur} \tag{3.22}$$

Since $f(\cdot)$ is a strictly increasing function (Lemma 3.1 [iii]), (3.22) implies that C_w^{ur} is a strictly decreasing function of Γ_w^u, while (3.21) implies that Γ_w^u is a strictly decreasing function of C_m^{ru}. Hence, we can derive a well-defined strictly increasing function such that $C_w^{ur} = h_1(C_m^{ru})$ from the equation system (3.17), (3.18) and (3.20). Additionally, it is worth noting that Γ_w^u reaches its lower bound when C_m^{ru} reaches its upper bound. If we substitute the lower bound of Γ_w^u, which equals I_m^u/I_w^u into (3.21), we derive the same maximum $\bar{C}_m^{ru}$ as defined in (3.11). Therefore, we conclude that for all (C_m^{ru}, C_w^{ur}) which are subject to $C_w^{ur} = h_1(C_m^{ru})$ and $C_m^{ru} < \bar{C}_m^{ru}$, only men migrate from rural to urban area, while women are indifferent between moving or not in equilibrium. Now it is straightforward to show that women are discouraged from migrating (from urban to rural area) if we further relax the condition such that $C_w^{ur} \geq h_1(C_m^{ru})$.

Next, let us consider the situation where only women migrate from urban to rural area while men are indifferent between moving or not. In this scenario, (3.17) and (3.18) still hold while (3.19) and (3.20) should be altered as follows:

$$\Delta I_m = \frac{-\Gamma_w^r I_m^u + \Gamma_w^u \Gamma_w^r I_w^u - \Gamma_w^u I_m^r + \Gamma_w^u \Gamma_w^r I_w^r}{\Gamma_w^r - \Gamma_w^u t} = 0 \tag{3.23}$$

$$\Delta I_w = \frac{I_m^u - \Gamma_w^u I_w^u + I_m^r - \Gamma_w^r I_w^r}{\Gamma_w^r - \Gamma_w^u} > 0$$

By using the same logic as in the previous arguments, we are able to derive a strictly increasing function, i.e. $C_m^{ru} = h_2(C_w^{ur})$ $(C_w^{ur} < \bar{C}_w^{ur})$ from the equation system (3.17), (3.18) and (3.23). Similarly, it can be concluded that for $C_m^{ru} \geq h_2(C_w^{ur})$ and $C_w^{ur} < \bar{C}_w^{ur}$ only women migrate from urban to rural area, while men stay.

We can summarise the results so far by establishing Corollary 3.3 as follows.

Corollary 3.3 *[i] $C_w^{ur} \geq h_1(C_m^{ru})$ and $C_m^{ru} = \bar{C}_m^{ru}$ are necessary conditions for the equilibrium where **only** men migrate from rural to urban area;*

*[ii] $C_m^{ru} \geq h_2(C_w^{ur})$ and $C_w^{ur} < \bar{C}_w^{ur}$ are necessary conditions for the equilibrium where **only***

women migrate from urban to rural area.

To see it more clearly, $C^{w}_{ur} = h_1(C^{m}_{ru})$ is derived from (3.17), (3.18) and (3.20) while $C^{ru}_{m} = h_2(C^{ur}_{w})$ from (3.17), (3.18) and (3.23). The only difference between these two equation systems lies in the constraints affected by whether it is ΔI_m or ΔI_w that equals zero. Furthermore, it is relatively simple to show that both $C^{ur}_{w} = h_1(C^{ru}_{m})$ and $C^{ru}_{m} = h_2(C^{ur}_{w})$ go through origin and $(\bar{C}^{ru}_{m}, \bar{C}^{ur}_{w})$[8].

However, so far the only information missing is whether $C^{ur}_{w} = h_1(C^{ru}_{m})$ is above or below $C^{ur}_{w} < h_2^{-1}(C^{ru}_{m})$, or they are intersected with each other. To explain what the boundaries of parameter space actually look like, we establish the following corollary.

Corollary 3.4 *[i] The curve $C^{ur}_{w} = h_1(C^{ru}_{m})$ is below the curve $C^{ur}_{w} < h_2^{-1}(C^{ru}_{m})$.*

*[ii] $h_1(C^{ru}_{m}) < C^{ur}_{w} < h_2^{-1}(C^{ru}_{m})$ and $C^{ru}_{m} < \bar{C}^{ru}_{m}$ are necessary conditions for the equilibrium where **both** sides of the markets are on the move.*

Proof. See Appendix. ■

From Corollary 3.2 to Corollary 3.4, we have exhausted all the scenarios that might arise, thus, all these conditions are both necessary and sufficient for the corresponding equilibrium/equilibria. We can sum up the parameter space and describe the corresponding equilibrium/equilibria in Figure 3.1.

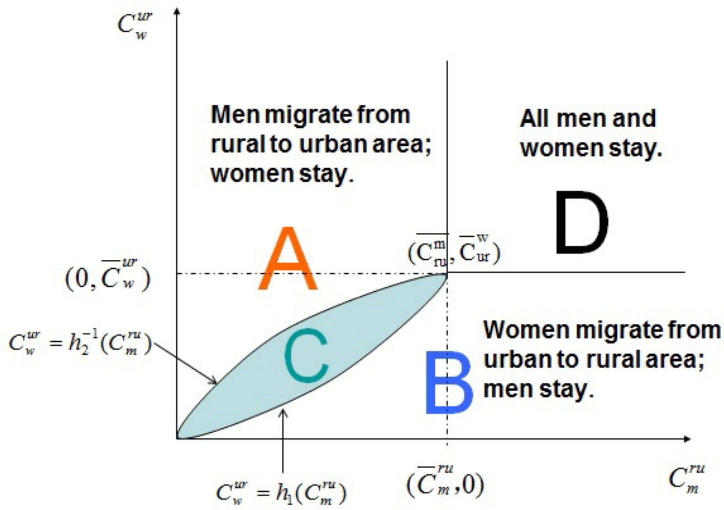

Figure 3.1: General description of the equilibrium set when $I^u_m/I^u_w < I^r_m/I^r_w$

[8]It can be verified simply by substituting the coordinates of origin and $(\bar{C}^{m}_{ru}, \bar{C}^{w}_{ur})$ into (3.17) and (3.18).

Summary 3.1 *In region A, i.e.* $C_w^{ur} \geq h_2^{-1}(C_m^{ru})$ *and* $C_m^{ru} < \bar{C}_m^{ru}$*, there is a unique equilibrium where only men migrate from rural to urban area.*

In region B, i.e. $C_m^{ru} \geq h_1^{-1}(C_w^{ur})$ *and* $C_w^{ur} < \bar{C}_w^{ur}$*, there is a unique equilibrium where only women migrate from urban to rural area.*

In region C, there are multiple equilibria, i.e. men migrate from rural to urban area, and/or women migrate from urban to rural area.

In region D, i.e. $C_m^{ru} \geq \bar{C}_m^{ru}$ *and* $C_w^{ur} \geq \bar{C}_w^{ur}$*, there is a unique equilibrium where neither side of the markets moves.*

The result we have derived is supported by some empirical evidence.

As we mentioned before, many of the Asian countries (i.e. China, Korea, Vietnam, etc.) have the rooted traditions of son preference. Thus, the sex ratios of the newly–born boys to girls are much higher in rural areas than in urban areas, which fits into our presumption in the model that $I_m^u/I_w^u < I_m^r/I_w^r$. These countries more often than not have relatively isolated rural and urban economies. Also, the relevant cost of migrating from urban to rural area is much higher than the other way around.[9] Thus, the fact that the majority of the workers migrating from rural to urban area are male can also be easily derived by using our model here (i.e. region A in Figure 3.1 in our model).

As the description of the equilibria in the above discussion is a bit abstract, we introduce a few numerical examples by assigning reasonable values to the parameters.

Example 3.1 *Given meeting efficiency* $\alpha_0 = 0.1$*, market leaving rate* $\delta = 0.02$*, discount rate* $\rho = 0.02$ *and the initial inflow* $I_m^u = 1.2$*,* $I_w^u = 1.0$*,* $I_m^r = 1.8$*,* $I_w^r = 1.0$*, calculate the unmatched population of the marriage market in equilibrium in the parameter space of* (C_m^{ru}, C_{ur}^w)*.*

First of all, we note in this example that $I_m^r/I_w^r > I_m^u/I_w^u$.

Then, we need to identify the pair of threshold migrating costs $(\bar{C}_m^{ru}, \bar{C}_{ur}^w)$ with which both men and women are indifferent between migrating or not. Thus, the *ultimate* inflows are the same as *initial* inflows. That is

$$\begin{aligned} \Gamma_w^u &= I_m^u/I_w^u = 1.2 \\ \Gamma_w^r &= I_m^r/I_w^r = 1.8 \end{aligned}$$

By using Lemma 3.1 (iii) and the expressions (3.11) and (3.12), the pair of threshold migrating costs can be solved as

$$\bar{C}_w^{ur} = 0.0586$$

[9] Here, the cost does not necessarily mean the physical cost, but rather the psychological barriers towards a different market, or it can be interpreted as the opportunity costs. For instance, urban areas generally have more convenient facilities for living in almost all aspects, thus, it is easier for the rural-born male to adapt to living in urban areas, but not vice versa.

$$\bar{C}_m^{ru} = 0.1762$$

Thus, neither side of the markets moves in equilibrium when the migrating costs exceed the threshold levels, i.e. in region D: $C_m^{ru} \geq \bar{C}_m^{ru} = 0.1762$ and $C_w^{ur} \geq \bar{C}_w^{ur} = 0.0586$.

The curves, i.e. $C_w^{ur} = h_1(C_m^{ru})$ and $C_w^{ur} = h_2^{-1}(C_m^{ru})$, can be derived by solving the equation systems, i.e. (3.17)(3.18)(3.20) and (3.17)(3.18)(3.23), respectively, with all the information provided.

Figure 3.2 illustrates how the unmatched population S_j^p responds to the change of C_m^{ru} within region A (i.e. $C_w^{ur} \geq h_2^{-1}(C_m^{ru})$ and $C_m^{ru} < \bar{C}_m^{ru} = 0.1762$). In this case, there are only men migrating from rural to urban area in equilibrium. We note from Figure 3.2 that S_m^u and S_m^r, as well as S_w^u and S_w^r converge when C_m^{ru} approaches zero, which coincides with our presumption in the example that $I_w^u = I_w^r$.

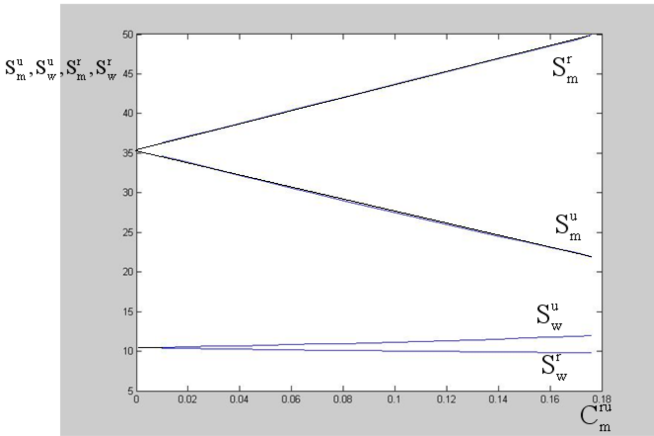

Figure 3.2: The unmatched populations S_j^p in equilibrium when $C_{ur}^w \geq h_2^{-1}(C_{ru}^m)$ and $C_{ru}^m <$ 0.1762

Similarly, when the prior migrating costs are in region B (i.e. $C_m^{ru} \geq h_1^{-1}(C_w^{ur})$ and $C_w^{ur} < \bar{C}_w^{ur} = 0.0586$), there are only women migrating from urban to rural area in equilibrium. Figure 3.3 illustrates that how the unmatched populations S_j^p respond to the change of C_w^{ur} in this case. We note from Figure 3.3 that, as C_w^{ur} approaches zero, S_m^u (S_w^u) is still less than S_m^r (S_w^r), which coincides with our presumption in the example that $I_m^u < I_m^r$.

As stated in Corollary 3.4, there are multiple equilibria where men migrate from rural to urban area and/or women migrate from urban to rural area when the prior migrating costs are in region C (i.e. $h_1(C_m^{ru}) < C_w^{ur} < h_2^{-1}(C_m^{ru})$ and $C_m^{ru} < \bar{C}_m^{ru} = 0.1762$ in this example). The equilibria where either side of the markets moves have already been covered in the last two cases; moreover, there is another equilibrium where both sides of the markets are on the move. For example, given. $C_m^{ru} = 0.08$ and $C_w^{ur} = 0.0242$, which are both inside region C, the migrating populations and the unmatched populations S_j^p in equilibrium can be derived by

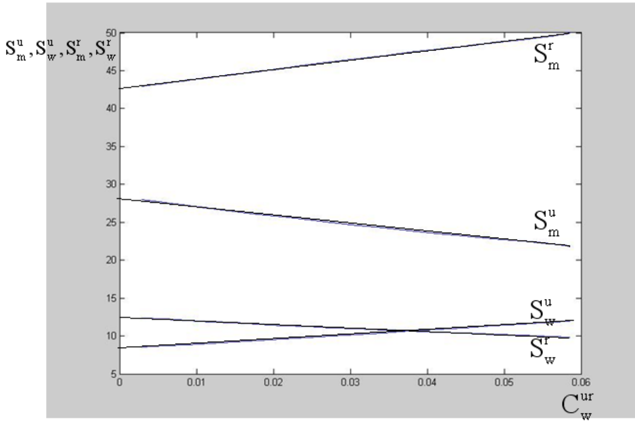

Figure 3.3: The unmatched populations S_j^p in equilibrium when $C_{ru}^m \geq h_1^{-1}(C_{ur}^w)$ and $C_{ur}^w < 0.0586$

solving the equation system (3.17)(3.18)(3.20)(3.23). That is, $\Delta I_m^{ru} = 0.0531$, $\Delta I_w^{ur} = 0.0739$, and $S_m^u = 26.5500$, $S_w^u = 10.1992$, $S_m^r = 44.5225$, $S_w^r = 10.8733$.

Additionally, recalling that we have $I_m^u/I_w^u = 1.2/1.8 < 1 = I_m^r/I_w^r$ in the example, we can see from both Figure 3.2 and Figure 3.3 that S_m^r and S_w^u are decreasing, while S_m^u and S_w^r are increasing in equilibrium as C_m^{ru} decreases. More specifically, the reduction in C_m^{ru} directly increases S_m^u and decreases S_m^r, and then decreases S_w^u and increases S_w^r through positive externalities. This result also can be applied when C_w^{ur} decreases.

In the next example, we would like to see how the unmatched populations S_j^p respond to the change of one *initial* inflow in equilibrium holding other parameters constant.

Example 3.2 *Given meeting efficiency* $\alpha_0 = 0.1$*, market leaving rate* $\delta = 0.02$*, discount rate* $\rho = 0.02$*, the cost for men migrating from rural area to urban area* $C_m^{ru} = 0.1055$*, the cost for women of migrating from urban to rural area* $C_w^{ur} = 0.0410$*, and the initial inflows* $I_m^u = 1.2$*,* $I_w^u = 1.0$*,* $I_w^r = 1.0$*; we let* I_m^r *vary from* 1.2 *to* 2.4 *in order to see how the unmatched populations* S_j^p *change in equilibrium.*

The equilibrium unmatched populations of the marriage markets are displayed in Figure 3.4. We observe from the figure that $S_m^u(I_m^r)$, $S_w^u(I_m^r)$, $S_m^r(I_m^r)$ and $S_w^r(I_m^r)$ all have kinks at their intersects with the dash line. On the left of the dash line, the market reaches an equilibrium that neither side of the markets moves; and on the right of the dash line, the market reaches an equilibrium where only men migrate from rural to urban area.

3.3.1 The Equilibrium Set with a Common Migrating Cost

In this subsection, we describe the equilibrium set in a two dimensional diagram where the horizontal axis represents the common migrating cost, i.e. $C_m^{ru} = C_w^{ur} = C$, and the vertical

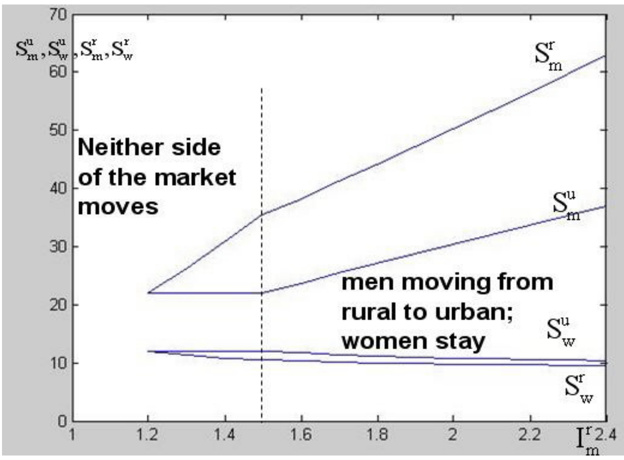

Figure 3.4: The equilibrium unmatched populations S_j^p when I_m^r varies from 1.2 to 2.4

axis represents the ratio of the urban and rural sex ratio of the *initial* inflows, i.e. $\frac{I_m^u/I_w^u}{I_m^r/I_w^r} = r < 1$.

By using Lemma 3.1 (iii), when men are indifferent between the two marriage markets the following condition must be satisfied:

$$f(I_w^u/I_m^u) - f(I_w^r/I_m^r) = f(\frac{1}{I_m^u/I_w^u}) - f(\frac{r}{I_m^u/I_w^u}) = C \text{ (Curve for men)} \tag{3.24}$$

Accordingly, when women are indifferent between the two marriage markets the condition is:

$$f(I_m^r/I_w^r) - f(I_m^u/I_w^u) = f(\frac{I_m^u/I_w^u}{r}) - f(I_m^u/I_w^u) = C \text{ (Curve for women)} \tag{3.25}$$

Since $f(\cdot)$ is an increasing function, the curves for men and women are downward sloping. The intuition here is that, as the gap between the urban and rural sex ratio of the *initial* inflows widens, it increases the incentive for men/women to migrate; thus, it requires higher migration cost to make them indifferent between the two markets. It is worth noting that the equation system (3.24) and (3.25) always has trivial solution such that $C = 0$ and $r = 1$. It is intuitive that, as the same ratio yields the same level of expected payoff for any new entrant, the migration cost becomes irrelevant.

Assume the sex ratio I_m^u/I_w^u is fixed. Also, recall that $f(\cdot)$ is increasing from zero and bounded above by $\dfrac{\alpha_0\theta}{1+\alpha_0\theta}$, then when r approaches zero we have:

$$\begin{aligned} f(I_w^u/I_m^u) - f(I_w^r/I_m^r) &= f(\frac{1}{I_m^u/I_w^u}) \\ f(I_m^r/I_w^r) - f(I_m^u/I_w^u) &= \frac{\alpha_0\theta}{1+\alpha_0\theta} - f(I_m^u/I_w^u) \end{aligned}$$

which are the horizontal intercepts for the two curves.

If the equation system only has the trivial solution, the two curves can be plotted in Figure 3.5 (It means that the two curves do not intersect for $C > 0$):

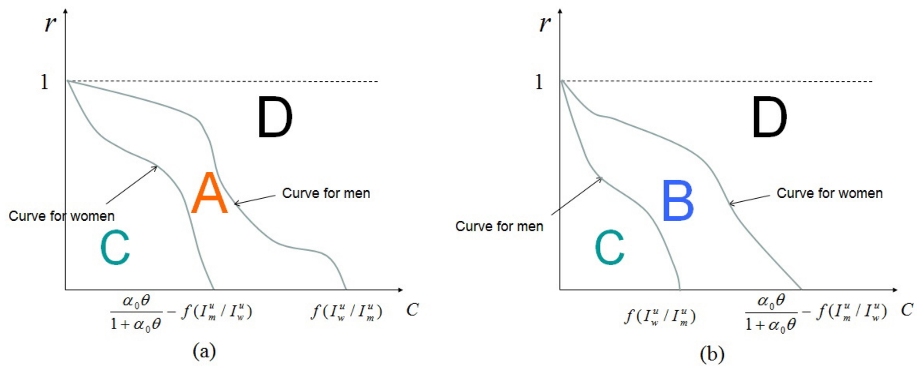

Figure 3.5: Curves for men and women do not intersect (a) when $f(\frac{1}{I^u_m/I^u_w}) > \dfrac{\alpha_0\theta}{1+\alpha_0\theta} - f(I^u_m/I^u_w)$; and (b) when $f(\frac{1}{I^u_m/I^u_w}) < \dfrac{\alpha_0\theta}{1+\alpha_0\theta} - f(I^u_m/I^u_w)$.

It is relatively easy to understand that, for parameters in region D in Figure 3.5 (when the cost of migration is high for both sexes), neither men or women migrate in equilibrium (which corresponds to region D in Figure 3.1). Especially, when the sex ratios in both markets are the same, i.e. $r = 1$ (the horizontal dashed line), people surely have no incentive to move.

For parameters in region A in Figure 3.5 (a) (when the cost of migration is too high for women but not for men), only men migrate in equilibrium (which corresponds to region A in Figure 3.1). Similarly, for parameters in region B in Figure 3.5 (b) (when the cost of migration is too high for men but not for women), only women migrate in equilibrium (which corresponds to region B in Figure 3.1).

For parameters in region C in Figure 3.5 (when the cost of migration is affordable for both sexes), there are multiple equilibria, i.e. men migrate from rural to urban area, and/or women migrate from urban to rural area (which corresponds to region C in Figure 3.1).

Next, we look at the case when the equation system (3.24) and (3.25) has nontrivial solution, the two curves can be plotted in Figure 3.6 (It means that the two curves intersect with each other for $C > 0$):

Generally speaking, the description of the equilibrium/equilibria for the parameters in region C and region D in Figure 3.6 are the same as in the previous case. In the region where the curve for women is on the right (left) of the curve for men, there are only men (women) migrating in equilibrium. And this results also hold for the case of having multiple intersections for the two curves.

At first glance, the descriptions of the equilibrium set in Figure 3.5 and Figure 3.6 here are very close to that in Figure 3.1 in the previous subsection, since all the diagrams divide the parameter space into three or four parts, each of which corresponds to a certain type of equilibrium/equilibria, i.e. in region A (B) only men (women) migrate, and in region D no one migrates, and in region C men migrate from rural to urban area and/or women migrate from

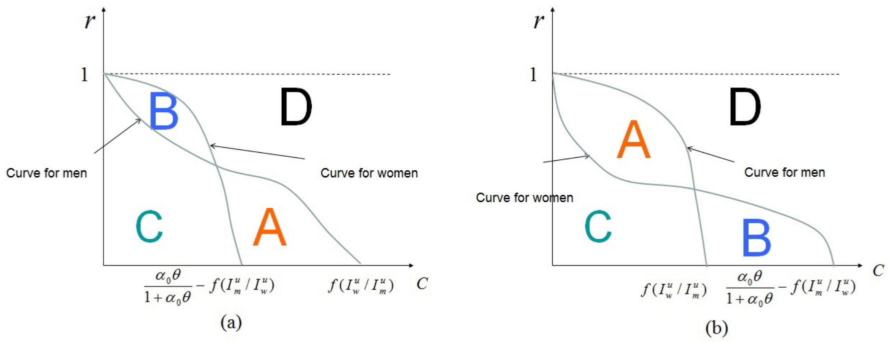

Figure 3.6: Curves for men and women intersect (a) when $f(\frac{1}{I^u_m/I^u_w}) > \dfrac{\alpha_0\theta}{1+\alpha_0\theta} - f(I^u_m/I^u_w)$; and (b) when $f(\frac{1}{I^u_m/I^u_w}) < \dfrac{\alpha_0\theta}{1+\alpha_0\theta} - f(I^u_m/I^u_w)$.

urban to rural area in equilibrium. The benefit is obvious. Figure 3.5 and Figure 3.6 allow us to link the contribution of the migration cost and the ratio of the urban and rural sex ratio of the initial inflows; and it helps us to gain more intuitions about how the type of equilibrium is jointly decided by the two factors.

It is relatively easy to comprehend that, when the migration cost is high enough, both sexes prefer to stay no matter how different the sex ratios in the two markets are. It is also not difficult to understand that, as the migration cost decreases and becomes affordable to *one* sex but not to the other, those who can afford cost would migrate to the other market in equilibrium. It follows that, as the migration cost becomes affordable to *both* sexes, both men and women *might* migrate to a different marriage market in equilibrium. However, it is worth noting that, as migration of men (women) from rural (urban) to urban (rural) area would bring positive externalities to the women (men) in urban (rural) area, the one-way migration equilibrium (either men or women migrate in equilibrium) might also rise in addition to the two-way migration equilibrium. As a matter of fact, it is a coordination problem of migration between rural men and urban women then.

3.4 Welfare Analysis

In this section, we conduct a welfare analysis to check whether welfare can be increased by subsidising towards migrating costs. Recall that we have assumed that each individual's utility derived from marriage equals the quality of her/his partner, which is normalised as unity. In this sense, we can sum up the expected present values of all *entering* agents to measure welfare.

Here, we still assume $I^u_m/I^u_w < I^r_m/I^r_w$. Suppose the relevant migrating costs are in region A, we learn from Summary 3.1 that in equilibrium there are only men migrating from rural to urban area, i.e. $\Delta I^{ru}_m > 0$ and $\Delta I^{ur}_w = 0$. On the one hand, if the government only partially subsidises women to reduce their cost of migrating to rural area within region A (i.e. C^{ur}_w moves down until it touches $C_w^{ur} = h_2^{-1}(C_m^{ru})$), the equilibrium remains the same and the policy is ineffective in improving the welfare. On the other hand, if the government subsidises men in order to reduce their cost of migrating to urban area within region A (i.e. C_m^{ru} moves leftward), the sex ratios of both urban and rural marriage market will converge. Thus, we mainly discuss how the welfare responds to the subsidies for only male migrants in this section.

Recalling that W_j^p denotes the expected present value of an unmatched agent of sex j still searching in marriage market p, it is straightforward to write the welfare function as follows.

$$\Pi = (W_m^u I_m'^u - C_m^{ru} \Delta I_m^{ru}) + W_w^u I_w'^u + W_m^r I_m'^r + W_w^r I_w'^r$$

We note in equilibrium that the migrating population (i.e. the rural-born men) must be indifferent between moving or staying; that is, $W_m^u - W_m^r = C_m^{ru}$. Thus, we are able to derive an important welfare function, which we can see in a more general form by using the *initial* inflows I_j^p and the sex ratios of the *ultimate* inflows Γ_j^p. That is

$$\begin{aligned} \Pi &= W_m^u I_m^u + W_w^u I_w^u + W_m^r I_m^r + W_w^r I_w^r \\ &= f(\Gamma_m^u) I_m^u + f(\Gamma_w^u) I_w^u + f(\Gamma_m^r) I_m^r + f(\Gamma_w^r) I_w^r \end{aligned} \tag{3.26}$$

Recalling that $I_m^u / I_w^u < I_m^r / I_w^r$ without loss of generality, we normalise the *initial* inflow of urban women as unity, whilst the other *initial* inflows can be written as

$$I_m^u = k_1,\ I_w^u = 1,\ I_m^r = k_2\lambda,\ I_w^r = \lambda \tag{3.27}$$

$$(I_m^u / I_w^u < I_m^r / I_w^r \text{ implies that } k_2 > k_1)$$

The sex ratios of the *ultimate* inflows can be written as

$$\Gamma_w^u = I_m'^u / I_w'^u = (I_m^u + \Delta I_m^{ru}) / I_w^u = k_1 + \Delta I_m^{ru} \tag{3.28}$$

$$\Gamma_w^r = I_m'^r / I_w'^r = (I_m^r - \Delta I_m^{ru}) / I_w^r = k_2 - \Delta I_m^{ru} / \lambda \tag{3.29}$$

Substituting (3.27) and (3.28) and (3.29) into the welfare function (3.26), we arrive at another form of the welfare function, expressed in terms of ΔI_m^{ru}. That is

$$\Pi = f(1/(k_1 + \Delta I_m^{ru}))k_1 + f(k_1 + \Delta I_m^{ru}) + f(1/(k_2 - \frac{\Delta I_m^{ru}}{\lambda}))k_2\lambda + f(k_2 - \frac{\Delta I_m^{ru}}{\lambda})\lambda$$

We differentiate the above welfare function with respect to the migrating cost C_m^{ru} and obtain

$$
\begin{aligned}
\frac{d\Pi}{dC_m^{ru}} &= [-\frac{k_1}{(k_1+\Delta I_m^{ru})^2}f'(\frac{1}{k_1+\Delta I_m^{ru}})+f'(k_1+\Delta I_m^{ru}) \\
&+\frac{k_2}{(k_2-\frac{\Delta I_m^{ru}}{\lambda})^2}f'(\frac{1}{k_2-\frac{\Delta I_m^{ru}}{\lambda}})-f'(k_2-\frac{\Delta I_m^{ru}}{\lambda})]\frac{d\Delta I_m^{ru}}{dC_m^{ru}}
\end{aligned}
$$

We need to note that $\frac{d\Delta I_m^{ru}}{dC_m^{ru}} < 0$ as the migrant population decreases in the migrating cost.

The main argument here is that the reduced migrating cost encourages more people (i.e. sex j agents) to move, which has positive (negative) externalities to the rest of the sex j agents in the birthplace (target) market and the sex $-j$ agents in the target (birthplace) market. But the aggregate effect of the externalities seems ambiguous so far. Instead, we decompose it into: [i] the male-side market welfare, which is $f(\Gamma_m^u)I_m^u + f(\Gamma_m^r)I_m^r$; and [ii] the female-side market welfare, which is $f(\Gamma_w^u)I_w^u + f(\Gamma_w^r)I_w^r$.

In order to simplify the analysis of comparative statics, we restrict ourselves to the case that the government *fully* subsidises the male migrants, which we denote by a subscript fs. In this case, the sex ratios of the *ultimate* inflows are same in equilibrium. That is

$$
\Gamma_{w,fs}^u = \Gamma_{w,fs}^r = \frac{k_1+k_2\lambda}{1+\lambda} \tag{3.30}
$$

Lemma 3.2 *When Γ_w^p is in the area where $f(\cdot)$ is concave (convex), the* ***full*** *subsidies for the male migrants increase (decrease) the female-side market welfare.*

Proof. When Γ_j^p is in the area where $f(\cdot)$ is concave (convex), it yields

$$
\begin{aligned}
& \frac{f(\Gamma_w^u)}{1+\lambda}+\frac{\lambda f(\Gamma_w^r)}{1+\lambda} \\
= & \frac{f(I_m'^u/I_w'^u)}{1+\lambda}+\frac{\lambda f(I_m'^r/I_w'^r)}{1+\lambda} \\
= & \frac{f(I_m'^u)}{1+\lambda}+\frac{\lambda f(I_m'^r/\lambda)}{1+\lambda} \\
\lesseqgtr & f(\frac{I_m'^u}{1+\lambda}+\frac{I_m'^r}{1+\lambda}) = f(\frac{k_1+\lambda k_2}{1+\lambda}) = f(\Gamma_{w,fs}^u) = f(\Gamma_{w,fs}^r)
\end{aligned}
$$

the last line is derived by using (3.30)

Multiplying both sides of the above inequality by $(1+\lambda)$, we obtain

$$
f(\Gamma_w^u)+\lambda f(\Gamma_w^r) \lesseqgtr f(\Gamma_{w,fs}^u)+\lambda f(\Gamma_{w,fs}^r)
$$

It follows that

$$
f(\Gamma_w^u)I_w^u + f(\Gamma_w^r)I_w^r \lesseqgtr f(\Gamma_{w,fs}^u)I_w^u + f(\Gamma_{w,fs}^r)I_w^r
$$

Here we complete the proof of Lemma 3.2. ∎

In Lemma 3.1 we show that $f(\cdot)$ is an increasing function without giving its explicit form.

We learn from Lemma 3.2 that the welfare is closely related to the concavity (convexity) of $f(\cdot)$. Figure 3.7 gives a general description of the shape of $f(\cdot)$ curve, which is convex (concave) when $\Gamma \lesseqgtr \Gamma^*$. It can be shown that Γ^* is always less than unity[10].

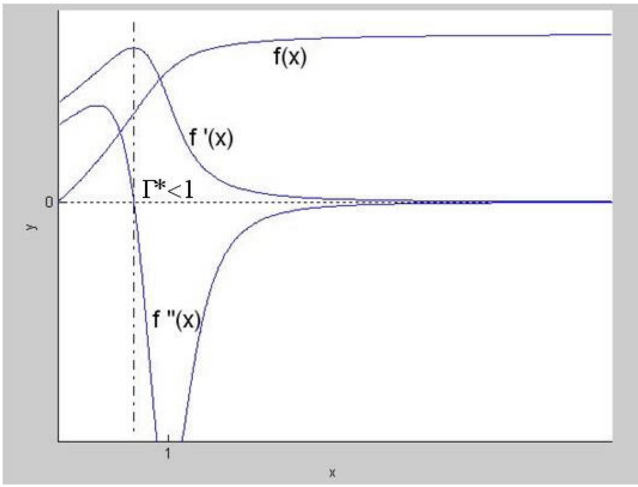

Figure 3.7: The shape of $f(x)$ curve

As we mentioned, whilst the migration of sex j brings positive (negative) externalities to the opposite sex $-j$ in the target (birthplace) market, it is unclear whether the aggregate welfare of sex $-j$ is affected. We establish the following corollary to provide the necessary condition under which the full subsidies toward one group would benefit the other group as a whole.

Corollary 3.5 *The full subsidies towards sex j migrants always increase the welfare of the opposite sex (i.e. sex $-j$) if sex $-j$ is the short side of each marriage market (i.e. $I_j^r/I_{-j}^r > 1$ and $I_j^u/I_{-j}^u > 1$).*

Proof. Suppose the short side of the marriage markets is female, then the settings in (3.27) fit into the conditions.

Recall that $I_m^r/I_w^r > 1$ and $I_m^u/I_w^u > 1$, which necessarily yields $\Gamma_w^u > 1$ and $\Gamma_w^r > 1$ in equilibrium. Thus, Γ_w^u and Γ_w^r are both located on the concave part of $f(\cdot)$ curve (Figure 3.7). Then, Lemma 3.2 implies that a full subsidy towards male migrants always increases the female-side welfare.

Here we complete the proof of Corollary 3.5. ■

Next, we introduce a numerical example to illustrate Lemma 3.2 and Corollary 3.5.

Example 3.3 *Given the meeting efficiency $\alpha_0 = 0.1$, the market leaving rate $\delta = 0.02$, the discount rate $\rho = 0.02$, the initial inflow $I_m^u = 1.5$, $I_w^u = 1.0$, $I_m^r = 1.9$, $I_w^r = 1.0$ and the*

[10] This statement is supported by running a number of tests, but we have not provided the details of the proof of this statement in this chapter.

migrating costs $C_m^{ru} = 0.08$*,* $C_w^{ur} = 0.04$*, compare the welfare before and after the men are fully subsidised for the migrating cost.*

Following the same calculation as in Example 1, the pair of threshold migrating costs can be derived as follows:

$$\begin{aligned} C_m^{ru} &= 0.0888 \\ C_w^{ur} &= 0.0213 \end{aligned}$$

We note that the prior costs $C_m^{ru} = 0.08$ and $C_w^{ur} = 0.04$ are in region A in Figure 3.1, which implies that there is a unique equilibrium where only men migrate from rural to urban area.

If the government fully subsidises the male migrants, the relevant sizes of the migrating population and the *ultimate* inflows, as well as the male-side (female-side) welfare, can be derived in Table 3.1:

	C_m^{ru}	C_w^{ur}	ΔI_m^{ru}	$I_m'^u$	$I_w'^u$	$I_m'^r$	$I_w'^r$	Π_m	Π_w
without subsidy	0.08	0.04	0.0192	1.5192	1	1.8808	1	1.0484	1.3306
full subsidy to men	0.00	0.04	0.2000	1.7000	1	1.7000	1	1.0469	1.3344

Table 3.1: The comparison of the male-(female-)side welfare with/without subsidy

Substituting $\alpha_0 = 0.1$, $\delta = 0.02$ and $\rho = 0.02$ into (3.9) and (3.10), we derive the inflexion $\Gamma^* = 0.6887$ in this specific example, i.e. $f(\cdot)$ is convex (concave) when $\Gamma \lesseqgtr 0.6887$.

We note from last column in Table 3.1 that, conditional on Γ_w^p being on the concave proportion of $f(\cdot)$ curve (i.e. $\Gamma_w^r = 1.8808 > 0.6887$ and $\Gamma_w^u = 1.5192 > 0.6887$), the female-side welfare is increased as the government subsidises men who are migrating to urban area. In contrast, we note from the second last column in the table that, conditional on Γ_m^p is on the convex proportion of $f(\cdot)$ curve (i.e. $\Gamma_m^r = 1/1.8808 < 0.6887$ and $\Gamma_m^u = 1/1.5192 < 0.6887$), the male-side welfare decreases even if it is men that receive a full subsidy.

All the results in the numerical example coincide with the statements in Lemma 3.2 and Corollary 3.5.

3.5 Conclusion

This chapter develops a two-sided matching model of marriage markets with homogeneous agents and non-transferable utility. We extend the framework of the one-market model by Burdett and Coles (1997) into a two-market one, and also incorporate the migrating cost which might alter the decisions of new entrants. While it is easy to understand that migration between markets can be the result of the different sex ratios, which are determined by the unbalanced inflows and meeting efficiency, it is important to note that the type of equilibrium (equilibria) also depend(s) on the relevant migrating costs. We have shown in the chapter that, with any fixed inflows of unbalanced sex ratios the markets always come to an equilibrium (equilibria), and the migrating directions can not be same for both groups of agents, if any (Proposition 3.2). That is, if migration turns out to be too costly for both sides of the markets, the markets reach an equilibrium where neither side of the market moves; if the migration costs are too high for one side but affordable for the other, only the latter migrate in equilibrium. However, if the migrating costs are equally affordable to both sides of the markets, there are multiple equilibria where either or both sides are on the move. The main idea here is that the migration of any agent brings positive (negative) externalities to the agents of same sex (the opposite sex) in the birthplace market and the opposite sex (the agents of same sex) in the target market. Besides, based on the assumptions of the much higher urban-to-rural-area migrating cost than vice versa, plus the higher sex-ratio-at-birth (boys-to-girls) in rural area, our model helps to explain the important force driving a large population of male migrants (rather than female migrants) into the urban cities from rural areas. Furthermore, we have characterised the equilibrium set in the two dimensional diagram where the horizontal axis represents the common migrating cost and the vertical axis represents the ratio of the urban and rural sex ratio.

We have explored the welfare implication within the framework of this model. The welfare is calibrated in terms of the expected payoffs of all entering agents. Thus it is a function of the sizes of each inflow and the sex ratios in each market. Instead of looking at the aggregate welfare of the inflows, we examine how the welfare of a single market-side responds to the subsidy of the migrating costs for the other side. Our analysis has shown that the full subsidy of the migrating costs does not necessarily benefit those who receive them but always benefits the opposite sex if they are the short sides of both markets.

We have not analysed such interesting issues as which exact equilibrium we'd be in if the migrating costs varied for different paths, since the multiple equilibria conclusion is not always a satisfying answer to real life problems. However, we might ignore this issue since the set of "equally affordable" migrating costs is relatively small, and in practice it is not likely for the cost parameter to fall in this narrow a set.

In the simple settings of our model, we have not considered the market of heterogeneous agents, which might better fit into the conditions that truly affect individuals' decisions when seeking a long-term partner. It is without a doubt a more complicated topic since it would inevitably involve the quality distribution of the unmatched population.

3.6 Appendix

Proof of Proposition 3.1. In SE, the inflows exactly replicate the outflows. Here, we only need to equate the volume of flows since the homogeneous market condition ensures a constant degenerated quality distribution. Given the meeting technology, the instantaneous number of successful matches on either side of market is $\alpha_0 S_m S_w/(S_m + S_w)$, which must be equal to the inflow, (if adding up the destruction of population due to physical reason). The conditions can be shown by establishing the following equation system:

$$\frac{\alpha_0 S_m S_w}{S_w + S_m} + \delta(S_w - \frac{\alpha_0 S_m S_w}{S_w + S_m}) = (1-\delta)\frac{\alpha_0 S_m S_w}{S_w + S_m} + \delta S_w = I_w \tag{3.31}$$

$$\frac{\alpha_0 S_m S_w}{S_w + S_m} + \delta(S_m - \frac{\alpha_0 S_m S_w}{S_w + S_m}) = (1-\delta)\frac{\alpha_0 S_m S_w}{S_w + S_m} + \delta S_m = I_m \tag{3.32}$$

Subtracting equation (3.31) from (3.32), we obtain

$$\delta(S_m - S_w) = I_m - I_w \Rightarrow S_m = S_w + \frac{I_m - I_w}{\delta} \tag{3.33}$$

Substituting (3.33) into (3.31), we obtain a quadratic equation in terms of S_w. That is

$$k_1 S_w^2 + k_2 S_w + k_3 = 0 \tag{3.34}$$

$$\text{where } k_1 = 2\delta^2 + \alpha_0\delta - \alpha_0\delta^2$$

$$k_2 = (\alpha_0 - \alpha_0\delta + \delta)(I_m - I_w) - 2\delta I_w$$

$$k_3 = -I_w(I_m - I_w)$$

The following statements show the uniqueness of the pair of positive solutions to equation (3.33) and (3.34).

Since $\delta \in (0,1)$, we have $k_1 = 2\delta^2 + \alpha_0\delta - \alpha_0\delta^2 > 0$; additionally, the discriminant is positive, which is

$$\begin{aligned} D &= k_2^2 - 4k_1k_3 \\ &= (\alpha_0 I_m - \alpha_0\delta I_m + \delta I_m + \alpha_0\delta I_w - \alpha_0 I_w - 3\delta I_w)^2 + 4I_w(I_m - I_w)(2\delta^2 + \alpha_0\delta - \alpha_0\delta^2) \\ &= (I_m - I_w)^2(\alpha_0 - \alpha_0\delta + \delta)^2 + 4I_m I_w\delta^2 > 0 \end{aligned}$$

If we assume without loss of generality that $I_m > I_w$, then it immediately yields that $k_1 > 0, k_3 < 0$ and $D > 0$, which imply only one positive root S_w^* to the quadratic equation (3.34). Thus, we obtain a unique pair of root (S_w^*, S_m^*) by substituting S_w^* into (3.33).

Alternatively, we can give out the unique pair of positive solutions mathematically. The roots of equation (3.34) are

$$S_w^* = \frac{2\delta I_w - (I_m - I_w)(\alpha_0 - \alpha_0\delta + \delta) \pm \sqrt{D}}{2\delta(\alpha_0 - \alpha_0\delta + 2\delta)}$$

The case of $I_m > I_w$ has been discussed already and proves to have one unique pair of positive roots to the simultaneous equations (3.31) and (3.32). Next, we move on to the case of $I_m < I_w$, which yields that $k_2 < 0$. By using basic algebra, $k_1 > 0$, $k_2 < 0$, $k_3 > 0$ and $D > 0$ imply that S^*_{w1} and S^*_{w2} are both positive. The first pair of roots are

$$S^*_{w1} = \frac{2\delta I_w - (I_m - I_w)(\alpha_0 - \alpha_0\delta + \delta) + \sqrt{D}}{2\delta(\alpha_0 - \alpha_0\delta + 2\delta)} > 0$$

$$\begin{aligned} S^*_{m1} &= \frac{2\delta I_w - (I_m - I_w)(\alpha_0 - \alpha_0\delta + \delta) + \sqrt{D}}{2\delta(\alpha_0 - \alpha_0\delta + 2\delta)} + \frac{I_m - I_w}{\delta} \\ &= \frac{2\delta I_w - (I_w - I_m)(\alpha_0 - \alpha_0\delta + \delta) + \sqrt{D}}{2\delta(\alpha_0 - \alpha_0\delta + 2\delta)} \\ &> \frac{2\delta I_w - (I_w - I_m)(\alpha_0 - \alpha_0\delta + \delta) + (I_w - I_m)(\alpha_0 - \alpha_0\delta + \delta)}{2\delta(\alpha_0 - \alpha_0\delta + 2\delta)} > 0 \end{aligned}$$

Accordingly, the second pair of roots are

$$S^*_{w2} = \frac{2\delta I_w - (I_m - I_w)(\alpha_0 - \alpha_0\delta + \delta) - \sqrt{D}}{2\delta(\alpha_0 - \alpha_0\delta + 2\delta)} > 0$$

$$\begin{aligned} S^*_{m2} &= \frac{2\delta I_w - (I_m - I_w)(\alpha_0 - \alpha_0\delta + \delta) - \sqrt{D}}{2\delta(\alpha_0 - \alpha_0\delta + 2\delta)} + \frac{I_m - I_w}{\delta} \\ &= \frac{2\delta I_w - (I_w - I_m)(\alpha_0 - \alpha_0\delta + \delta) - \sqrt{D}}{2\delta(\alpha_0 - \alpha_0\delta + 2\delta)} \\ &< \frac{2\delta I_w - (I_w - I_m)\delta - \sqrt{(I_m - I_w)^2\delta^2 + 4I_m I_w \delta^2}}{2\delta(\alpha_0 - \alpha_0\delta + 2\delta)} \\ &= \frac{\delta(I_w + I_m) - \sqrt{(I_m + I_w)^2\delta^2}}{2\delta(\alpha_0 - \alpha_0\delta + 2\delta)} = 0 \end{aligned}$$

We note that the second pair of roots should be ruled out by a on-negativity constraint.

It is worth noting that when $\delta = 0$ there is no root for the equation system (3.31) and (3.32), thus $\delta > 0$ is a necessary condition for the equilibrium with unbalanced inflows.

Here, we have completed the proof of Proposition 3.1. ■

Proof of Corollary 3.4. As for the necessary conditions for the equilibrium that both sides of the market move, (3.17) and (3.18) shall still hold, but the constraints on ΔI_m and ΔI_w are relaxed as follows:

$$\Delta I_m = \frac{-\Gamma^r_w I^u_m + \Gamma^u_w \Gamma^r_w I^u_w - \Gamma^u_w I^r_m + \Gamma^u_w \Gamma^r_w I^r_w}{\Gamma^r_w - \Gamma^u_w} > 0 \tag{3.35}$$

$$\Delta I_w = \frac{I^u_m - \Gamma^u_w I^u_w + I^r_m - \Gamma^r_w I^r_w}{\Gamma^r_w - \Gamma^u_w} > 0 \tag{3.36}$$

Intuitively, since (3.17) and (3.18) consist an indeterminate equation system of C^{ur}_w and

C_m^{ru}, we can not derive C_w^{ur} exclusively in terms of C_m^{ru}. Instead, we try to find how C_w^{ur} ranges at each level of C_m^{ru}. In the following paragraphs, we intend to obtain the range of C_w^{ur} by allowing Γ_w^r and/or Γ_w^u to vary within their constraints, i.e. (3.35) and (3.36). However, in order to avoid further confusion, we divide the proof into four parts.

First of all, we derive the correlation between Γ_w^r and ΔI_w as follows:

Holding C_m^{ru} constant and differentiating both sides of (3.17) with respect of Γ_w^r, it yields

$$\frac{\partial f(1/\Gamma_w^u)}{\partial(1/\Gamma_w^u)}(-\frac{1}{\Gamma_w^{u2}})\frac{\partial \Gamma_w^u}{\partial \Gamma_w^r} - \frac{\partial f(1/\Gamma_w^r)}{\partial(1/\Gamma_w^r)}(-\frac{1}{\Gamma_w^{r2}}) = 0$$

Since $f(\cdot)$ is a strictly increasing function, we have $\frac{\partial \Gamma_w^u}{\partial \Gamma_w^r} > 0$.

Recall that $\dfrac{I_m^u + \Delta I_m}{I_w^u - \Delta I_w} = \dfrac{I_m'^u}{I_w'^u} = \Gamma_w^u$ and $\dfrac{I_m^r - \Delta I_m}{I_w^r + \Delta I_w} = \dfrac{I_m'^r}{I_w'^r} = \Gamma_w^r$. Substituting for ΔI_m, we obtain

$$\Gamma_w^u(I_w^u - \Delta I_w) - I_m^u = I_m^r - \Gamma_w^r(I_w^r + \Delta I_w) \tag{3.37}$$

Differentiating both sides of (3.37) with respect of ΔI_w, we obtain

$$\frac{\partial \Gamma_w^u}{\partial \Gamma_w^r}\frac{\partial \Gamma_w^r}{\partial \Delta I_w}(I_w^u - \Delta I_w) - \Gamma_w^u = -\frac{\partial \Gamma_w^r}{\partial \Delta I_w}(I_w^r + \Delta I_w) - \Gamma_w^r$$

$$\Rightarrow \frac{\partial \Gamma_w^r}{\partial \Delta I_w}[\frac{\partial \Gamma_w^u}{\partial \Gamma_w^r}(I_w^u - \Delta I_w) + (I_w^r + \Delta I_w)] = \Gamma_w^u - \Gamma_w^r \tag{3.38}$$

Since we have assumed $I_m^u/I_w^u < I_m^r/I_w^r$, we will have $\Gamma_w^u < \Gamma_w^r$ in equilibrium. Also, as $\frac{\partial \Gamma_w^u}{\partial \Gamma_w^r} > 0$, $I_w^u - \Delta I_w > 0$ and $I_w^r + \Delta I_w > 0$, (3.38) implies that

$$\frac{\partial \Gamma_w^r}{\partial \Delta I_w} < 0 \text{ (and similarly } \frac{\partial \Gamma_w^r}{\partial \Delta I_m} > 0) \tag{3.39}$$

Secondly, we derive the correlation between C_{ur}^w and ΔI_w as follows:

We differentiate (3.17) by Γ_w^r, and obtain

$$\begin{aligned}\frac{\partial W_m^u}{\partial \Gamma_w^r} - \frac{\partial W_m^r}{\partial \Gamma_w^r} &= \frac{\partial W_m^u}{\partial M_m^u}\frac{\partial M_m^u}{\partial \Gamma_w^r} - \frac{\partial W_m^r}{\partial M_m^r}\frac{\partial M_m^r}{\partial \Gamma_w^r} = \frac{\partial g(1/M_w^u)}{\partial \Gamma_w^r} - \frac{\partial g(1/M_w^r)}{\partial \Gamma_w^r} \\ &= \frac{\partial g(1/M_w^u)}{\partial(1/M_w^u)}\frac{\partial(1/M_w^u)}{\partial M_w^u}\frac{\partial M_w^u}{\partial \Gamma_w^r} - \frac{\partial g(1/M_w^r)}{\partial(M_w^r)}\frac{\partial(1/M_w^r)}{\partial M_w^r}\frac{\partial M_w^r}{\partial \Gamma_w^r} = 0 \end{aligned} \tag{3.40}$$

Recalling from the proof in Lemma 3.1 that $W_j^p = \dfrac{\alpha_0\theta}{1/M_j^p + 1 + \alpha_0\theta}$, we obtain

$$\frac{\partial W_j^p}{\partial M_j^p} = \frac{\alpha_0\theta}{(1 + M_j^p + \alpha_0\theta M_j^p)^2} \tag{3.41}$$

Substituting (3.41) into (3.40), it follows that

$$\frac{\partial W_m^u}{\partial \Gamma_w^r} - \frac{\partial W_m^r}{\partial \Gamma_w^r} = \frac{\alpha_0\theta}{(1 + 1/M_w^u + \alpha_0\theta/M_w^u)^2}(-\frac{1}{M_w^{u2}})\frac{\partial M_w^u}{\partial \Gamma_w^r}$$

$$-\frac{\alpha_0\theta}{(1+1/M_w^r+\alpha_0\theta/M_w^r)^2}(-\frac{1}{M_w^{r2}})\frac{\partial M_w^r}{\partial \Gamma_w^r}=0$$

$$\Rightarrow \frac{\partial M_w^u}{\partial \Gamma_w^r}=\frac{\partial M_w^r}{\partial \Gamma_w^r}\frac{(M_w^u+1+\alpha_0\theta)^2}{(M_w^r+1+\alpha_0\theta)^2}>0 \tag{3.42}$$

Thirdly, we derive the correlation between C_w^{ur} and Γ_w^r as follows:

Differentiating (3.18) by Γ_w^r, we obtain

$$\begin{aligned}
\frac{\partial C_w^{ur}}{\partial \Gamma_w^r} &= \frac{\partial W_w^r}{\partial \Gamma_w^r}-\frac{\partial W_w^u}{\partial \Gamma_w^r} \\
&= \frac{\partial W_w^r}{\partial M_w^r}\frac{\partial M_w^r}{\partial \Gamma_w^r}-\frac{\partial W_w^u}{\partial M_w^u}\frac{\partial M_w^u}{\partial \Gamma_w^r} \\
&= \frac{\alpha_0\theta}{(1+M_w^r+\alpha_0\theta M_w^r)^2}\frac{\partial M_w^r}{\partial \Gamma_w^r}-\frac{\alpha_0\theta}{(1+M_w^u+\alpha_0\theta M_w^u)^2}\frac{\partial M_w^u}{\partial \Gamma_w^r} \\
&\quad \text{(substitute for } \frac{\partial M_w^u}{\partial \Gamma_w^r} \text{ by using (3.42))} \\
&= \frac{\alpha_0\theta}{(1+M_w^r+\alpha_0\theta M_w^r)^2}\frac{\partial M_w^r}{\partial \Gamma_w^r}-\frac{\alpha_0\theta}{(1+M_w^u+\alpha_0\theta M_w^u)^2}\frac{\partial M_w^r}{\partial \Gamma_w}\frac{(M_w^u+1+\alpha_0\theta)^2}{(M_w^r+1+\alpha_0\theta)^2} \\
&= \alpha_0\theta\cdot\frac{\partial M_w^r}{\partial \Gamma_w^r}(\frac{1}{1+M_w^r+\alpha_0\theta M_w^r}+\frac{M_w^u+1+\alpha_0\theta}{M_w^r+1+\alpha_0\theta}) \\
&\quad \cdot(\frac{1}{1+M_w^r+\alpha_0\theta M_w^r}-\frac{M_w^u+1+\alpha_0\theta}{(1+M_w^u+\alpha_0\theta M_w^u)(M_w^r+1+\alpha_0\theta)}) \\
&= \alpha_0\theta\cdot\frac{\partial M_w^r}{\partial \Gamma_w^r}(\frac{1}{1+M_w^r+\alpha_0\theta M_w^r}+\frac{M_w^u+1+\alpha_0\theta}{M_w^r+1+\alpha_0\theta}) \\
&\quad \cdot\frac{(2+\alpha_0\theta)(M_w^u-M_w^r)}{(1+M_w^r+\alpha_0\theta M_w^r)(1+M_w^u+\alpha_0\theta M_w^u)(M_w^r+1+\alpha_0\theta)}
\end{aligned} \tag{3.43}$$

Since we have assumed $I_m^u/I_w^u < I_m^r/I_w^r$, it implies that $\Gamma_w^u < \Gamma_w^r$. Also, as we learn from the proof in Lemma 3.1 that M_w^p increases in Γ_w^p, it is easy to have $M_w^u < M_w^r$ and $\frac{\partial M_w^r}{\partial \Gamma_w^r} > 0$. Therefore, we can now conclude from (3.43) that

$$\frac{\partial C_w^{ur}}{\partial \Gamma_w^r}<0 \tag{3.44}$$

Finally, combining the results of (3.44) with (3.39), we arrive at the conclusion that C_{ur}^w is an increasing (decreasing) function of ΔI_w (ΔI_m). It implies that the lower (upper) bound of C_w^{ur} coincides with the lower bound of ΔI_w (ΔI_m) which equals zero. Recalling that $C_w^{ur}=h_1(C_m^{ru})$ is defined by the equation system (3.17) and (3.18) plus $\Delta I_w = 0$, while $C_m^{ru}=h_2(C_w^{ur})$ is defined by (3.17) and (3.18) plus $\Delta I_m = 0$, we arrive at the following inequality.

$$h_1(C_m^{ru})<C_w^{ur}<h_2^{-1}(C_m^{ru})$$

Now we have completed all the proofs in Corollary 3.4. ∎

Chapter 4

Wage Setting and Migration in Oligopsony with Recruiting Limits

Abstract

In this chapter the migration is of workers to jobs. We consider a spatial oligopsony model in which two co-locating small firms with recruiting capacity constraints and a large firm without such limit are competing for workers along a "strip" market. We discuss how the recruiting capacity affects the intra-group competition and hence the inter-group competition. We show that when the recruiting limits shrink the expected wages offered by the small firms decrease, while the wage offered by the big firm increases. This in turn helps to explain the recent trend of wage disparity between public and private jobs. We also provide a three-stage-game framework (the workers decide whether to relocate in the first stage, then the big firm decides its wage offer, and lastly, the two co-locating firms simultaneous set wages), and characterise the equilibrium wages and the size and direction of migration.

4.1 Introduction

We consider a spatial oligopsony model in which two co-locating small firms with recruiting capacity constraints and a big firm without such limit are competing for workers along a "strip" market. The two types of employers in the model can be many private firms that are constrained by the recruiting limits, and the public sector, which is relative flexible in size. Besides, the distribution of the workers on the Hotelling line indicates each individual's job preference, i.e. some might place a higher value on the public services, which brings them a sense of "warm glow" (Andreoni, 1990), whilst others may care more about the monetary payment. The wage differentials between the public and private sectors have been intensively explored in empirical research. Most literature finds that the workers in public services receive premium pay, which can mainly be attributed to the more pervasive unions (Smith, 1976). However, other studies find a negative differential (i.e. the workers in private sectors receive a pay premium, especially in developing economies) and attribute it to the efficiency wage provided by the private firms, the compensation of higher job security or the non-wage benefit in the public sector. In other words, the wage premium offered by the private firms are the joint result of inter-sector and intra-sector competition. In this paper, we show that the employers' recruiting capacities also constitute an important factor in this wage disparity. Among them, the recruiting limits of the small firms are more sensitive to the economic environment, for instance, in recession many employees suffer a salary cut or even unemployment, especially in the private sector. Meanwhile many workers in the public sector enjoy a pay rise. It is relatively easy to comprehend the former by the argument of the negative demand shock, however, there is limited economic explanation for the latter. Worse still in an economic recession, the general public become very frustrated by such headlines in newspapers that public sector pay races ahead of inflation. We show in our model that such an unusual trend of wage disparity between the two types of employers (i.e. the salary cut in the private sector together with the pay rise in the public sector) can be attributed to the shrinking recruiting capacities, which are themselves caused by a general anticipation of doom often experienced in economic downturn.

In the conventional study of oligopoly (oligopsony), most of the literature focuses on analysing the behaviour of oligopolists (oligopsonists) upon some prior determined elasticity of demand (supply) in a homogeneous market. For instance, in the product market, both Cournot (Quantity) Competition and Bertrand (Price) Competition explicitly assume a downward sloping market demand curve. Among them, the former requires an auctioneer who clears the market according to the demand schedule (i.e. setting an equilibrium price in accordance with the realised outputs by all producers), and the latter is pure price competition. Here, we intend to probe the mechanism of the labour market, and will take advantage of the latter model since it fits more realistically with the fact that in a labour market, firms often use wage as a strategic variable rather than depending on a fictional auctioneer to clear the market by setting an equilibrium wage.

However, there are three major limits of this model. Firstly, in the market where each consumer can only consume one unit of good (or each worker can only provide one unit of labour), the price elasticity of demand (or the wage elasticity of labour supply) will be infinite for the identical producers (employers), which is incompatible with the assumption of the above models. It makes more sense that, in the real-life circumstance, the price (wage) competition is imperfect and a price cut (or a wage cut) will not result in the immediate takeover of the whole market (or the resignation of all employees at once). Thus, we introduce heterogeneous preferences of workers for the jobs that make the labour supply more inelastic. Secondly, the conventional literature always skirts around the issue of the strategic behaviour of the firms, without elaborating on the interaction from the other side of the market. It is true that in the labour market, either the agents on one side of the market (i.e. workers) might need to overcome some physical or psychological cost to reach the other side of the market (i.e. employers), or they would choose to pay a large sum of cost to relocate to a new physical or psychological position which would enable them to enjoy higher (expected) utility. Thirdly, Bertrand Competition is often attacked on the grounds that many firms often have producing (or recruiting) limits that disable them from taking over the whole market. For example, this is typically the case for many private firms which are constrained by recruiting capacities, whilst the public sector is relatively flexible in size. In this case, if one firm does set the wage at the break-even level, the other firm can always enjoy the residual labour supply, which implies that the firms do not necessarily set the wage at the break-even level in equilibrium when there is capacity constraint.

A model of wage-setting duopsony is a natural starting point for analysing the behaviour of oligopsonists. Such a model has been fully solved under the assumption of unlimited recruiting capacity at a constant marginal revenue product of labour. In this case the unique equilibrium occurs where each firm sets the break-even wage (if the firms are equal in marginal revenue product of labour) and yields zero profit (i.e. Bertrand Equilibrium). We can consider the case of recruiting capacity for an employer as the extreme case of decreasing returns to scale. As we mentioned before, in the multi-firm case where the employers do not have adequate capacity to recruit all the workers in the market, they would set wages below the break-even level and enjoy positive profits (i.e. Bertrand-Edgeworth Model). With the development of studies on Industrial Organization, there emerges lots of literature on the price-setting game amongst oligopolists. As a matter of fact, oligopoly and oligopsony are mathematically equivalent, so the explicit discussion of the price-setting game with capacity constraints in the literature sheds much light on the behaviour of the oligopsonists in our study. For instance, the seminal paper by Kreps and Scheinkman (1983) models a two-stage game and proves that the Bertrand Competition with production precommitment is equivalent to the Cournot Competition in a duopoly model with a deterministic demand curve, but in our paper we assume the capacity is exogenous. Hirata (2009) solves explicitly the Bertrand-Edgeworth equilibrium in a three-firm oligopoly with asymmetric capacity, however, we assume symmetric capacity for the small employers and unlimited capacity for the big employer in our model, which we believe has more intuitive implications. The research paper by Bhaskar, Manning and To (2002) has a very detailed discussion of duopolists' wage-posting strategy in the horizontal differentiated labour market, as well as its implications in wage dispersion and minimum wages. We borrow the same idea of horizontal differentiation in labour and expand it into the oligopsonistic competition of three firms. Moreover, we assign the employers with limited and unlimited recruiting capacities in our model, which to the best of our knowledge is the first to address its implications in the trend of wage disparity in public and private sector by using such theoretical setup.

The structure of the paper can be divided into two parts: the wage-setting game and the three-stage game.

In the wage-setting game, our focus is on the wage-posting strategy by the employers and the horizontal heterogeneity of workers. We model a spatial market where three oligopsonists (i.e. two co-locating small firms with recruiting capacity constraints and a large firm without such limit) compete for a limited number of workers distributed along a "strip" market. For the moment, the setup of the spatial market as well as the "physical" distribution of the workers along it can be a metaphor of heterogeneous preferences or the ability of the individual workers to work for a specific firm. In Hotelling's (1929) and Salop's (1979) papers, their notion of the traveling costs to work can be literally interpreted as the actual commuting costs incurred (physical distance), or as a subjective measure of an individual's preference of one job over another (psychological distance), or as the training cost to offset the differentials in ability to a universal skill requirement (suitability distance). In any of the interpretations of the horizontal heterogeneity in this model, the worker is willing to "travel" to the further or "the less preferred" employer for a sufficient wage premium. In the following section, we characterise the equilibrium wages posted by all firms in the wage-setting game, given a uniform distribution of the workers along the "strip" market.

One novelty of our model is that it also brings in competition from a third firm, which itself does not have recruiting limit at the other end of the "strip" market. In the price-setting (wage-setting) game, the rationing rule of the residual demand (supply) affects the expected profit (production) of the firms. Therefore, we need to elucidate the rationing scheme before any conclusion of the consequent strategic behaviour can be discussed. In the literature on oligopoly, Beckmann (1965) explicitly calculates the equilibrium by constructing a symmetric Bertrand-Edgeworth model and using the *proportional* rationing rule (which means the high-priced firm's residual demand at any price is proportional to the overall market at that price). Levitan and Shubik (1972) obtain the equilibrium of the symmetric model by using the efficient rationing rule (which means the highest-value consumers go to the low-priced firm). In spite of the different rationing rules, both papers reach similar conclusions. That is: (i) the two firms act as if they can combine their capacities together (i.e. they would produce up to each of their capacities) and charge a universal monopolistic price (pure strategy) when their capacities are small; (ii) the two firms randomise their wage within an interval (an *atomless*, symmetric mixed strategy equilibrium) when their capacities grow larger; (iii) they set the break-even price when the capacities can meet the needs of the whole market. In this paper, we apply the *efficient* rationing rule and find a similar conclusion: (i) The small firms post a single but relatively low wage, as though they combine the recruiting capacity together (pure strategy) when their recruiting limits are low; (ii) they randomise their wages within an interval (i.e. a symmetric mixed strategy but not necessarily atomless) when capacities grow larger; (iii) the employer at the far end without capacity limit constantly adopts pure strategy in posting wage.

Apart from the elaboration of types of wage-posting strategies, we also find that the (expected) wages posted by the small employers are decreasing, whilst those posted by big the employer are increasing as the capacities of the former shrink. The logic here is that the shrinking capacity of the small firms lowers the wage competition between the two, and hence makes it less expensive for the big firm to recruit more workers through positive externalities.

As mentioned before, the wage-setting game can be applied in a more general context (i.e. the Hotelling preferences), while the second part of this paper (i.e. the three-stage game) is more specific as we interpret the Hotelling line as the real traveling space. Here, we take into account the possible interaction of the other side of the market, i.e. the relocation decision of the workers. We propose the following framework for the three-stage model: (i) in the first stage, the workers decide whether to relocate themselves, and of course permanent migration requires a fixed cost; (ii) observing the realised distribution of workers along the market, the big employer decides its wage offer; (iii) then, the two co-locating firms simultaneously and non-cooperatively post wages. The sequence of the three-stage game is in line with the fact that it's much easier for the employers to adjust the wage offers than the employees' response of relocation. In this paper we look at the extreme case that workers are initially located at the two ends of the market respectively. In the discussion, we characterise the equilibrium wages posted by firms as well as the size and direction of the migration.

In the next section we describe the model. In the first part of section 4.2, we discuss the wage-setting game and characterise the equilibrium strategies played by all firms; in the second part we apply the results of the former section in the three-stage-game framework, and characterise the equilibrium; and in the third part we put forward some implications of the previous models. In Section 4.3 we consider possible extensions and conclude.

4.2 The Model

Consider the case of having three oligopsonists. We assume that a unit measure of workers are located along the interval A-B with two co-locating small firms (with recruiting limits k)[1] at City A and a large firm (without recruiting limit) at City B. We only focus on the case when $k \leq \frac{1}{2}$ throughout the paper as we are more interested in the wage competition between the firms with very different sizes.

All the firms produce identical products. For the sake of simplicity, we assume that they all have identical production function $Q(L) = L$ where labour L is the only factor of production. Each firm is a price taker and all the goods produced can be sold at price P. Thus, the marginal revenue product of labour is equal to the marginal revenue, which is $MRP_L = MR = P$.

4.2.1 A Wage-setting Game

In this subsection, we assume that the workers are uniformly distributed along a straight, mile-long district between City A and City B. The heterogeneity in physical locations implies that the workers have different wages net of traveling costs incurred. All workers have the same reservation wage to stay employed (i.e. workers would like to accept all positive wages net of commuting cost). It implies a monopsony context where the labour supply curve is linear, such that $L(w) = \dfrac{w}{c}$, where c is the transportation cost per mile for the potential employees. Since $MRP_L = P$, all the firms will pay the workers no more than unit price of the production P. Furthermore, for the sake of simplicity, let us assume $c = P = 1$, which implies that the marginal worker is located right at the other city when the firm posts a wage up to the marginal revenue P.

Concerning the labour market, each firm simultaneously and noncooperatively posts a wage offer, w_{Ai} by firm i at City A (where $i = 1, 2$), and w_B by the firm at City B. It's a simultaneous game, so the basic idea of the model can be presented generally as having the following two aspects:

1) The two co-locating firms are bidding for the workers near City A. However, since they are constrained by recruiting limits, we draw an instant intuition from the short-run price competition in the Bertrand-Edgeworth Model that it generates either pure or mixed strategy equilibrium which depends on the size of recruiting constraints, and both firms receive positive profits.

2) If the firms in different cities are *close* enough, (here *close* means there is some overlap of the labour force if we put together separately the equilibrium results of different cities), they will compete for the workers in the middle region of the "strip" market.

Now let us consider *only* the wage competition between the two co-locating small firms at City A (that is, ignoring the potential competing effect of their counterpart, the large firm, in City B). It is easy to comprehend that in equilibrium both firms charge the competitive wage: $w_{Ai} = w_{Aj} = P$ and receive zero profit if they have infinite recruiting capacities (i.e. Bertrand Equilibrium). However, as we have assumed limited recruiting capacities for the small firms, the competitive wage is no longer the equilibrium. The intuition is: supposing one firm lowers its wage slightly, it can readily make positive profit by recruiting the residual workers since the

[1] The recruiting capacity constraint can be attributed to diseconomies of scale for the small firms.

other firm can not absorb all the workers due to its recruiting cap. Of course, before explicitly solving for the equilibrium, we need to specify in which manner the workers are rationed. Here, we assume *efficient* rationing rule when there is wage overbidding, which implies that the high-wage firm recruits the high-value workers who are located nearer[2]. That is, if firm i posts a higher wage than firm j, it will absorb all available workers up to its recruiting capacity. The firm that loses in wage competition can only recruit the residual workers, if any. We also assume that the two firms would equally split the labour supply when they post a common wage. To sum up, the labour supply for firm i can be written as follows.

$$L_{Ai}(w_{Ai}) = \begin{cases} \min(w_{Ai}, k) & \text{if} \quad w_{Ai} > w_{Aj} \\ \min(\frac{w_{Ai}}{2}, k) & \text{if} \quad w_{Ai} = w_{Aj} \\ \min\{\max(w_{Ai} - k, 0), k\} & \text{if} \quad w_{Ai} < w_{Aj} \end{cases} \tag{4.1}$$

Next, we take into account the wage competition from the large firm at the far end of the "strip" market (i.e. City B). But first let us look at a simpler case of having only one firm in City A, the wage competition between firms from different cities can be illustrated in Figure 4.1.

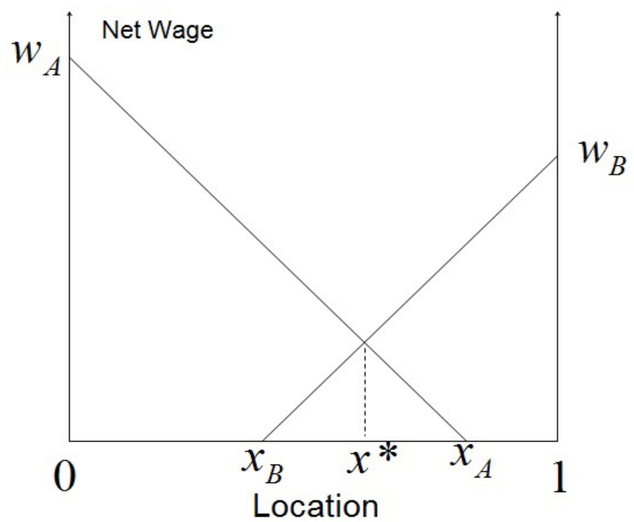

Figure 4.1: Illustration of workers' job choice with transportation costs

The horizontal axis represents the uniformly distributed workers along the mile-long "strip" market, and the vertical axis denotes the *net* wage received by workers. Recalling the per mile transportation cost $c = 1$, then a worker who is located x miles from City A would have to incur a cost of x if he/she is employed by the firm in City A, or to incur a cost of $(1 - x)$ if employed by the firm in City B. Now suppose that there is one firm at each end of the market, i.e. a firm in City A offers wage w_A and a firm in City B (which we will simply call Firm B from now on) offers w_B. Since we have assumed that the workers accept all positive wages net

[2]Here, higher value refers to the higher wage net of transportation cost, thus, the high-value workers are those who are located nearer to the firm. We consider the assumption of efficient rationing more reasonable, in that the workers who live nearer have better information about wage offers and can respond quicker to the firm that posts a higher wage, or the employers prefer to recruit "local" residents in spite of their identical productivity.

of commuting cost, it is straightforward that when the two net wage lines do not intersect, the firm in City A recruits the workers to the left of x_A while Firm B recruits the workers to the right of x_B, leaving those who are located between x_A and x_B unemployed. When the two net wage lines intersect, the workers to the left of the intersection x^* prefer to work in City A while those to the right of x^* prefer to work in City B. Therefore, we can see that the intersection of the net wage lines indicates the competition between the firms from different cities.

Let us look at the marginal worker who is indifferent between working in City A and in City B when the two net wage lines intersect. That is,

$$w_A - x^* = w_B - (1 - x^*) \Rightarrow x^* = \frac{w_A - w_B + 1}{2}$$

The labour supply for the firm in City A (which is denoted by L_A) is

$$L_A(w_A) = \frac{w_A - w_B + 1}{2} \tag{4.2}$$

Likewise, the labour supply for the firm in City B is

$$L_B(w_B) = \frac{w_B - w_A + 1}{2} \tag{4.3}$$

However, we need to keep in mind that the above expressions are for the case of having a single firm at each city, thus, an extension must be made to fit into the three-firm model. The labour supply function of the co-locating firm i in City A can be derived as

$$L_{Ai}(w_{Ai}) = \begin{cases} \min(w_{Ai}, k) & \text{if} \quad w_{Aj} < w_{Ai} \leq 1 - w_B \\ \min(\frac{w_{Ai} - w_B + 1}{2}, k) & \text{if} \quad 1 - w_B < w_{Ai} \text{ and } w_{Aj} < w_{Ai} \\ \min\{\max(w_{Ai} - k, 0), k\} & \text{if} \quad w_{Ai} \leq 1 - w_B \text{ and } w_{Ai} < w_{Aj} \\ \min\{\max(\frac{w_{Ai} - w_B + 1}{2} - k, 0), k\} & \text{if} \quad 1 - w_B < w_{Ai} < w_{Aj} \\ \min(\dfrac{w_{Ai}}{2}, k) & \text{if} \quad w_{Ai} = w_{Aj} \leq 1 - w_B \\ \min(\frac{w_{Ai} - w_B + 1}{4}, k) & \text{if} \quad 1 - w_B < w_{Ai} = w_{Aj} \end{cases} \tag{4.4}$$

Next, let us move on to Firm B which does not have recruiting limit. It is relatively simple to obtain its labour supply function as $L_B(w_B) = w_B$ if there is no intersection of the net wage lines (i.e. $w_B < 1 - \max(w_{Ai}, w_{Aj})$ and there is no competition between the firms coming from different ends). However, Firm B's labour supply function evolves to a more complicated pattern if the net wage lines intersect. It needs to be discussed in two cases. Before that, let us denote the intersect of the net wage line of Firm B with that of the higher-wage firm in City A by x_H^*, and the lower-wage firm by x_L^*.

Case 4.1 *The net wage line of firm B only intersects with that of the higher-wage firm in City A, i.e.* $1 - \max(w_{Ai}, w_{Aj}) < w_B \leq 1 - \min(w_{Ai}, w_{Aj})$. *(See Figure 4.2)*

In Case 4.1, when $k \geq x_H^*$ the higher-wage firm in City A employs the labour to the left of x_H^* and Firm B absorbs the rest, leaving the lower-wage firm in City A empty handed; when $1 - w_B < k < x_H^*$, the higher-wage firm in City A recruits workers to it capacity k leaving the rest to Firm B only; when $\min(w_{Ai}, w_{Aj}) < k \leq 1 - w_B$ the higher-wage firm in City A recruits to its capacity k but Firm B only has the workers to the right of $(1 - w_B)$, leaving the lower-wage firm in City A recruits no one; when $k \leq \min(w_{Ai}, w_{Aj})$, the higher-wage firm in

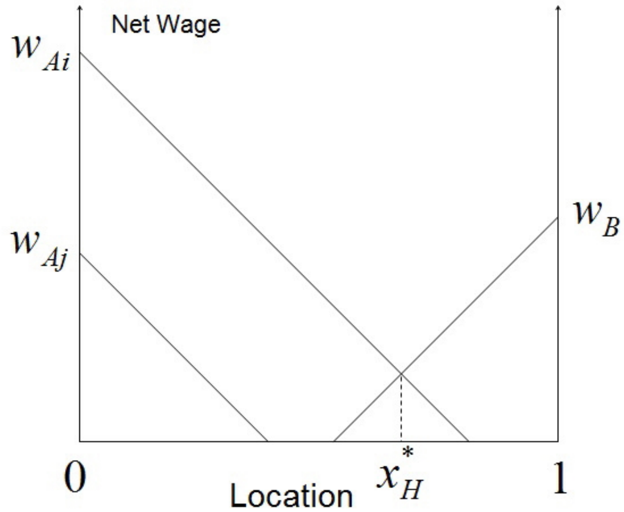

Figure 4.2: Illustration of the net wage lines in Case 4.1

City A recruits to its capacity k, and Firm B has the workers to the right of $(1 - w_B)$, and the lower-wage firm in City A recruits $\min(\min(w_{Ai}, w_{Aj}) - k, k)$.

Case 4.2 *The net wage line of Firm B intersects with the net wage lines of both firms in City A, i.e.* $w_B > 1 - \min(w_{Ai}, w_{Aj})$. *(See Figure 4.3)*

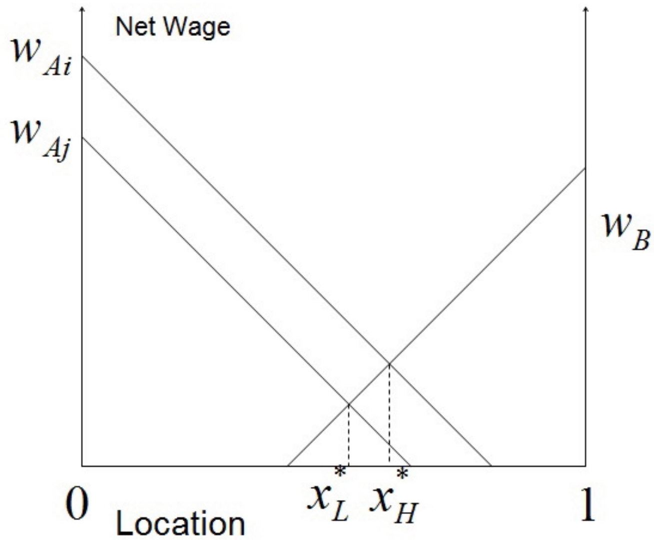

Figure 4.3: Illustration the net wage lines in Case 4.2

Similarly, when $k \geq x_H^*$ the higher-wage firm in City A employs the labour to the left of x_H^*, and Firm B absorbs the rest, leaving the lower-wage firm empty handed; when $x_L^* \leq k < x_H^*$ the higher-wage firm in City A recruits workers to its capacity k, leaving the rest to Firm B only; when $\frac{x_L^*}{2} \leq k < x_L^*$ the higher-wage firm in City A employs up to it capacity k, but Firm B only has the workers to the right of x_L^*, and the lower-wage firm in City A recruits

$(1 - x_L^* - k)$; when $k < \frac{x_L^*}{2}$ both firms in City A employ up to the capacity k and Firm B employs $\min(1-2k, 1-w_B)$.

To sum up, the labour supply function of the firm in City B can be written as follows:

1) NO intersect, i.e. $w_B \leq 1 - \max(w_{Ai}, w_{Aj})$	$L_B(w_B) = w_B$
2) ONE intersect (Case 4.1)	
i.e. $1 - \max(w_{Ai}, w_{Aj}) < w_B \leq 1 - \min(w_{Ai}, w_{Aj})$	
(i) and $k > \frac{\max(w_{Ai},w_{Aj})-w_B+1}{2}$	$L_B(w_B) = \frac{w_B-\max(w_{Ai},w_{Aj})+1}{2}$
(ii) and $1 - w_B \leq k \leq \frac{\max(w_{Ai},w_{Aj})-w_B+1}{2}$	$L_B(w_B) = 1 - k$
(iii) and $k \leq 1 - w_B$	$L_B(w_B) = w_B$
3) TWO intersects (Case 4.2), i.e. $w_B > 1 - \min(w_{Ai}, w_{Aj})$	
(i) and $k > \frac{\max(w_{Ai},w_{Aj})-w_B+1}{2}$	$L_B(w_B) = \frac{w_B-\max(w_{Ai},w_{Aj})+1}{2}$
(ii) and $\frac{\min(w_{Ai},w_{Aj})-w_B+1}{2} < k \leq \frac{\max(w_{Ai},w_{Aj})-w_B+1}{2}$	$L_B(w_B) = 1 - k$
(iii) and $k \leq \frac{\min(w_{Ai},w_{Aj})-w_B+1}{2}$	$L_B(w_B) = \frac{w_B-\min(w_{Ai},w_{Aj})+1}{2}$

So far, the labour supply for all firms have been outlined as above. Let us proceed to the discussion of the optimal strategies by these firms in equilibrium.

Proposition 4.1 *If and only if* $0 < k \leq \frac{1}{3}$, *there exists a unique pure strategy Nash Equilibrium that both firms in City A post a common wage offer of* $2k$ *and Firm B posts a wage offer of* $\min(1-2k, \frac{1}{2})$.

Proof. First, we show that no competition happens between the firms from different cities when $0 < k \leq \frac{1}{4}$.

We note that the aggregate labour demand for the firms in City A will not exceed $\frac{1}{2}$ due to their recruiting limits k. Also, as the profit function of Firm B is $\pi_B = w_B(1 - w_B)$, it is straightforward to show that it would maximise the profit by setting the wage as high as $\frac{1}{2}$ and recruits the workers to the right of $x_B = \frac{1}{2}$. Then, it can be implied that there is no overlap of competition for the firms from different cities. Thus, the profit function for the firms in City A can be written as

$$\pi_{Ai}(w_{Ai}) = \begin{cases} \min(w_{Ai}, k)(1 - w_{Ai}) & \text{if} \quad w_{Ai} > w_{Aj} \\ \min(\frac{w_{Ai}}{2}, k)(1 - w_{Ai}) & \text{if} \quad w_{Ai} = w_{Aj} \\ \min\{\max(w_{Ai} - k, 0), k\}(1 - w_{Ai}) & \text{if} \quad w_{Ai} < w_{Aj} \end{cases} \tag{4.5}$$

We start by looking at the symmetric case when $w_{Ai} = w_{Aj}$. If $k > \frac{w_{Ai}}{2}$, then the first order condition of the second line in (4.5) yields $\arg\max \pi_{Ai}(w_{Ai}) = \frac{1}{2}$ which contradicts $\frac{w_{Ai}}{2} < k \leq \frac{1}{4}$. If $k \leq \frac{w_{Ai}}{2}$, then the profit function becomes $\pi_{Ai}(w_{Ai}) = k(1 - w_{Ai})$ which is monotonically decreasing in w_{Ai}. It yields $\arg\max \pi_{Ai}(w_{Ai}) = 2k$ as both firms in City A recruit to their capacities and have no incentives to deviate up by posting higher wages. On the other hand, if one firm lowers its wage from $2k$ to $(2k - \epsilon)$ where ϵ is small and positive, then its profit changes from $k(1-2k)$ to $(2k-\epsilon-k)(1-2k+\epsilon)$. We note that the change in profit is negative for $0 < k \leq \frac{1}{3}$. Thus, in this way we have precluded the asymmetric strategies in equilibrium.

Next comes the case that $\frac{1}{4} < k \leq \frac{1}{3}$. We first look at the situation where there is the same marginal worker for the firms in City A and Firm B, which implies that the profit function for the former is the same as (4.5). Again, we look at the symmetric case first. If $k > \dfrac{w_{Ai}}{2}$ the first order condition of the second line in (4.5) yields $\arg\max \pi_{Ai}(w_{Ai}) = \frac{1}{2}$ and each firm absorbs $\frac{1}{4}$

workers. However, at that time, either firm would enjoy a jump in labour supply by overbidding the wage a little bit, and thus $w_{Ai} = w_{Aj} = \frac{1}{2}$ can not be an equilibrium strategy. If $k \leq \frac{w_{Ai}}{2}$, by using the same argument as in the previous paragraph, we can see that the firms in City A only choose to post a common wage of $2k$ in equilibrium and again the asymmetric strategies are ruled out.

Let us move on to the strategy of Firm B. When $\frac{1}{4} < k \leq \frac{1}{3}$, (as we have shown that its counterparts in City A adopt a symmetric strategy which is $w_{Ai} = w_{Aj} = 2k$,) the labour supply function for Firm B can be easily written as

$$\pi_B(w_B) = \begin{cases} w_B(1 - w_B) & \text{if} \quad w_B \leq 1 - 2k \\ \frac{(w_B - 2k + 1)(1 - w_B)}{2} & \text{if} \quad w_B > 1 - 2k \end{cases} \tag{4.6}$$

If $w_B > 1 - 2k$, the first order condition of the second line in (4.6) yields $\arg\max \pi_B(w_B) = k$ which contradicts $\frac{1-w_B}{2} < k \leq \frac{1}{3}$. If $w_B \leq 1 - 2k$, we obtain $w_B \leq 1 - 2k < \frac{1}{2}$ since $\frac{1}{4} < k \leq \frac{1}{3}$. Hence, the first line in (4.6) yields that $\arg\max \pi_B(w_B) = 1 - 2k$. It means that when $\frac{1}{4} < k \leq \frac{1}{3}$ and the firms in City A pay $2k$, the optimal strategy for Firm B is to pay $(1 - 2k)$ and not to compete with firms in City A.

Moreover, we are able to gain some intuition about equilibrium strategies when $k > \frac{1}{3}$. If both firms in City A adopt pure symmetric strategy (i.e. each firm in City A posts a common wage that is greater or equal to $2k$), then as shown in previous argument, either of them can benefit by deviating downward. And if they post a common wage such that $w_{Ai} = w_{Aj} < 2k$, either firm can benefit by overbidding. Therefore, we rule out the pure strategy for the firms in City A when $k > \frac{1}{3}$. ■

The intuition of Proposition 4.1 is that when $0 < k \leq \frac{1}{3}$ the small employers always post wage at the joint monopolistic level as if there is no pressure of competition from the far end. Moreover, the proof in Proposition 4.1 also implies that, even without considering the possible competition from Firm B, the two firms in City A will adopt mixed strategy in posting wages when $k > \frac{1}{3}$. Moreover, we learn from the above proof that Firm B would benefit by raising the wage when $k > \frac{1}{3}$. Thus, we conjecture that there is a unique equilibrium when k exceeds $\frac{1}{3}$, the firms in City A adopt mixed strategy and Firm B adopts pure strategy.

Recall that mixed strategy equilibrium is to follow a pair of probability distribution over the respective strategy spaces, and each participant will choose any strategy with positive probability in order to maximise its own expected payoffs against its counterpart's strategy mix. Therefore, the top priority for solving the mixed strategy lies in the identification of the lower and upper bound of the wage offers posted by firms in City A. We denote the lower and upper bound of the mixed strategy by $\underline{w}_A$ and $\bar{w}_A$ respectively, and denote the c.d.f. by $G(w_A)$.

Next, we introduce a corollary to state the overlap of the strategies for firms in different cities, and it will be followed by a proposition that characterises the equilibrium strategies for all firms when k exceeds $\frac{1}{3}$.

Corollary 4.1 *For $\frac{1}{3} < k \leq \frac{1}{2}$, the pure strategy for Firm B overlaps with the whole boundaries of the mixed strategy for the firms in City A; that is $1 - w_B \leq \underline{w}_A \leq \bar{w}_A$.*

Proof. We prove $1 - w_B \leq \underline{w}_A$ by contradiction.

Suppose $1 - w_B > \underline{w}_A$ holds. If a firm picks the bottom wage $\underline{w}_A$ which is surely below the wage offered by its local counterpart, we can write the profit function by using the third line in

(4.4). Also, in mixed strategy equilibrium all the participants have *positive* expected payoffs, and thus the profit function can be written as

$$\pi(\underline{w}_A) = (\underline{w}_A - k)(1 - \underline{w}_A)$$

Profit maximisation implies that $\underline{w}_A = \frac{1+k}{2}$.

Recall that $\frac{1}{3} < k \leq \frac{1}{2}$, thus we have $\underline{w}_A \in (\frac{2}{3}, \frac{3}{4}]$. As we learn from the proof in Proposition (4.1) that $w_B > \frac{1}{3}$, it contradicts our presumption that $1 - w_B > \underline{w}_A$. Thus, it is straightforward to yield the results in Corollary 4.1. ■

Proposition 4.2 *For* $\frac{1}{3} < k \leq \frac{1}{2}$

(i) the firms in City A randomise their wage offers within the interval $[k + \frac{1}{2}, \frac{13k - 2k^2 - 2}{9k}]$ *following a c.d.f. of* $G(t) = \frac{(2k-3t+1)^2}{3(t-1)(10k-3t-1)}$*;*

(ii) Firm B adopts pure strategy in wage posting, i.e. $w_B = \frac{2(1-k)}{3}$.

Proof. We first derive the profit function by making the firm in City A choose the bottom wage. Corollary 4.1 implies that the fourth line of (4.4) is the labour supply for the firm. Thus, given the wage w_B posted by Firm B, the profit function can be written as follows.

$$\pi_A(\underline{w}_A) = (\frac{\underline{w}_A - w_B + 1}{2} - k)(1 - \underline{w}_A) \tag{4.7}$$

Profit maximisation implies:

$$\underline{w}_A = \frac{w_B}{2} + k \tag{4.8}$$

Substituting for $\underline{w}_A$ in (4.7) by using (4.8), we obtain the expected profit for the firm in City A, which is $\frac{(2-w_B-2k)^2}{8}$.

Next, consider the situation where the firm picks the top wage, then it will surely offer a higher wage than its local counterpart. We also learn from (4.7) that $\frac{\bar{w}_A - w_B + 1}{2} > k$, which implies that the second line in (4.4) is the labour supply for the firm. Thus, the profit function can be written as

$$\pi_A(\bar{w}_A) = k(1 - \bar{w}_A) \tag{4.9}$$

One feature of mixed strategy equilibrium is that each player is indifferent among all the actions that he or she selects, conditional on the mixed strategy of his/her counterparts. It implies that the expected payoffs (profits) of the firm are the same whenever it picks lowest or highest wage. Thus, (4.9) is equal to the expected profit we have derived previously. That is

$$\pi_A(\bar{w}_A) = k(1 - \bar{w}_A) = \frac{(2 - w_B - 2k)^2}{8} \tag{4.10}$$

Solving for $\bar{w}_A$, we obtain

$$\bar{w}_A = 1 - \frac{(2 - w_B - 2k)^2}{8k} \tag{4.11}$$

The general profit function for the firm in City A can be derived as

$$\pi_A(w_A) = G(w_A)k(1 - w_A) + [1 - G(w_A)](\frac{w_A - w_B + 1}{2} - k)(1 - w_A) \tag{4.12}$$

where $G(\cdot)$is the c.d.f. for the mixed strategy

As we mentioned before, in mixed strategy equilibrium each player has the same payoffs among all his/her actions, again we can equate (4.12) to the expected payoffs $\frac{(2-w_B-2k)^2}{8}$. Thus, we have derived the c.d.f. for the mixed strategy for the firms in City A, given the wage posted by Firm B. That is

$$G(w_A, w_B) = \frac{(2w_A - w_B - 2k)^2}{4\,(1-w_A)\,(w_B + 4k - w_A - 1)} \tag{4.13}$$

So far, the mixed strategy (which is also the best response) for the firms in City A has been properly characterised in (4.8) (4.11) and (4.13).

Next, let's turn to the wage-posting strategy by Firm B.

We learn from Corollary 4.1 that the net wage line of Firm B intersects with the net wage lines of both firms in City A. Also, we learn from (4.7) that $\frac{w_A - w_B + 1}{2} \geq k$. Thus, the labour supply function (iii) in Case 2 should be used to derive the profit for Firm B. That is

$$\pi_B(w_B) = \frac{w_B - \min(w_{A,i}, w_{A,j}) + 1}{2}(1 - w_B)$$

Since the firms in City A randomly choose wages, $\min(w_{A,i}, w_{A,j})$ (i.e. the first order statistic) is a random variable which follows a c.d.f. of $[1-(1-G)^2]$. Thus, the *expected* profit function for Firm B is

$$\begin{aligned} E\pi_B(w_B) &= (1-w_B)\int_{w_B/2+k}^{1-(2-w_B-2k)^2/8k} \frac{w_B - t + 1}{2} d[1-(1-G(t))^2] \\ &= (1-w_B)[\frac{w_B}{2} + \frac{(2-w_B-2k)^2}{16k} + \frac{1}{2}\int_{w_B/2+k}^{1-(2-w_B-2k)^2/8k} (2G(t) - G^2(t))dt] \\ &= \frac{(1-w_B)(w_B+2-2k)}{4} - \frac{(1-w_B)}{2}\int_{w_B/2+k}^{1-(2-w_B-2k)^2/8k} (1-G(t))^2 dt \end{aligned} \tag{4.14}$$

In order to find the maximised expected profit for Firm B, we examine the first derivative of (4.14), which is

$$\begin{aligned} \frac{dE\pi_B(w_B)}{dw_B} &= w_B[\int_{w_B/2+k}^{1-(2-w_B-2k)^2/8k} (1-G(t))^2 dt - 1]/2 + (k-1)/2 \\ &\quad -(1-w_B)\frac{d\int_{w_B/2+k}^{1-(2-w_B-2k)^2/8k}[1-G(t)]^2 dt}{dw_B} \end{aligned}$$

It can be checked that $\frac{dE\pi_B(w_B)}{dw_B} < 0$[3]. Also recall from Corollary 4.1 that w_B has a lower bound of $(1 - \underline{w}_A)$, thus, we obtain

$$\begin{aligned} \arg\max_{w_B} E\pi_B(w_B) &= 1 - (\frac{w_B}{2} + k) \\ &\Rightarrow \; w_B = \frac{2(1-k)}{3} \end{aligned} \tag{4.15}$$

[3] We have not included the analytical proof of this statement, but a number of numerical examples have been carried out to show a negative derivative.

Substituting for w_B in (4.8) (4.11) and (4.13) by using (4.15), we obtain

$$\begin{aligned} \underline{w}_A &= \frac{w_B}{2} + k = \frac{2k+1}{3} \\ \bar{w}_A &= \frac{13k - 2k^2 - 2}{9k} \\ G(w_A) &= \frac{(2k - 3w_A + 1)^2}{3(1 - w_A)(10k - 3w_A - 1)} \end{aligned}$$

So far, the mixed strategy for the firms in City A and the pure strategy for Firm B are fully characterised by the above expressions, and we have completed all the proofs of Proposition 4.2. ■

To have a better idea how the equilibrium strategies for all firms look like, we introduce the following example to illustrate the equilibrium wage(s) posted at a specific level of recruiting capacity. It will be followed by a more comprehensive graph which shows the comparative statics with respect to k.

Example 4.1 *Suppose each firm in City A has a recruiting capacity of $k = 0.4$. Calculate the equilibrium wage(s) posted by the firms in City A and Firm B.*

Since $k = 0.4 > \frac{1}{3}$, as stated in Proposition 4.2, the firms in City A adopt a mixed strategy to post their wage and Firm B uses pure strategy. Solving for w_B, $\underline{w}_A$, $\bar{w}_A$ and $G(w_A)$, we obtain that the firms in City A randomise their wage offers within $[0.6, 0.8]$ which follows a c.d.f. of $G(t) = (t - 0.6)^2 / (t - 1)^2$, and Firm B adopts the pure strategy of paying 0.4. We depict the strategies played by the firms in Figure 4.4 and Figure 4.5.

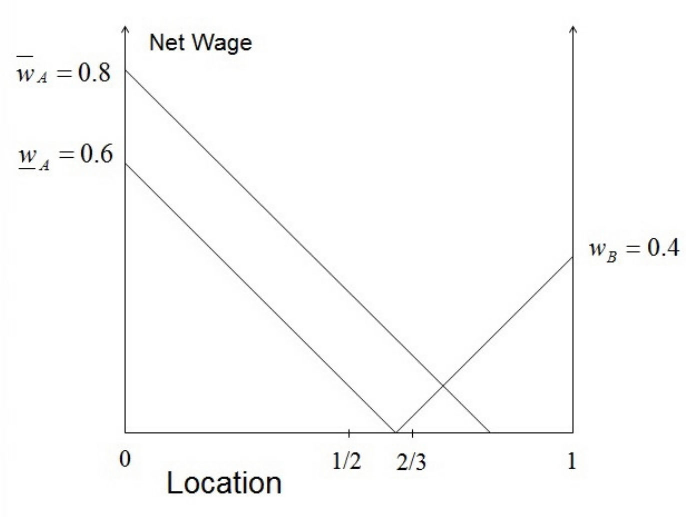

Figure 4.4: The boundary of the mixed strategy for the firms in City A and the pure strategy for Firm B

So far, the discussion of the parameter space of recruiting capacity k is complete. In order to further understand the wage-setting mechanism in the current model, we plot all the equilibrium wage(s) with respect to each specific level of recruiting limit k in Figure 4.6.

When $0 < k \leq \frac{1}{3}$, each firm in City A posts a common wage of $2k$ and recruit to capacity (i.e. the dark solid line till $k = \frac{1}{3}$) while Firm B posts a wage of $\min(\frac{1}{2}, 1 - 2k)$ (i.e. the red solid lines with a kink at $k = \frac{1}{4}$) (Proposition 4.1).

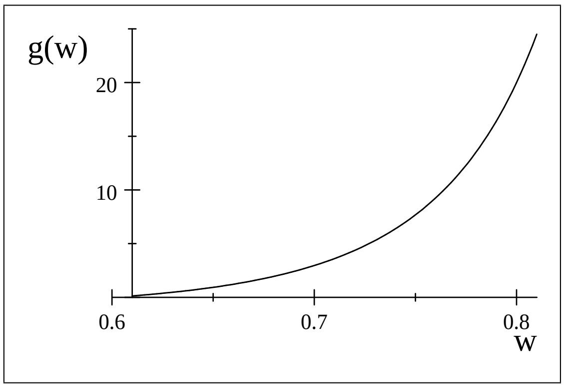

Figure 4.5: p.d.f. of the mixed strategy played by the firms in City A.

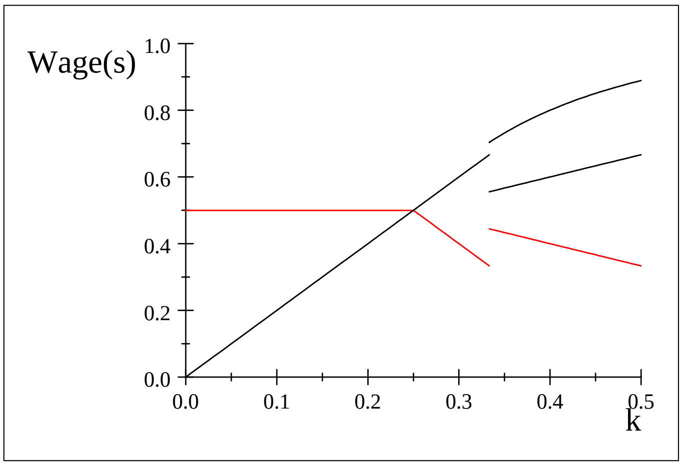

Figure 4.6: Equilibrium strategies of all firms with respect to the recruiting limits k

When $\frac{1}{3} < k \leq \frac{1}{2}$, taking into account the competition from Firm B, the firms in City A randomise wage offers within the lower and upper bound (i.e. the dark solid curves to the right of $k = 1/3$), while Firm B posts a wage which is represented by red line to the right of $k = 1/3$ (Proposition 4.2).

We observe form Figure 4.6 that when the recruiting limit exceeds the critical level of $\frac{1}{3}$ there is a sudden jump in the strategies of all firms. The intuition is that, suppose Firm B still uses the pure strategy by posting a wage of $(1-2k)$. On the one hand, if it raises wage from $(1-2k)$ to $(1-2k+\epsilon)$, the profit would change from $(1-2k)2k$ to $\frac{1-2k+\epsilon-2k+1}{2}(2k-\epsilon)$ for ϵ is small and positive. We note that the change is negative for $0 < k \leq \frac{1}{3}$, and positive for $\frac{1}{3} \leq k \leq \frac{1}{2}$. On the other hand, if it lowers the wage from $(1-2k)$ to $(1-2k-\epsilon)$, the profit would change from $(1-2k)2k$ to $(1-2k-\epsilon)(2k+\epsilon)$. We note that the change is negative for $k > \frac{1}{4}$, and positive for $k \leq \frac{1}{4}$. Thus, it implies when k slightly exceeds $\frac{1}{3}$ Firm B will have the incentive to deviate upward from $(1-2k) = \frac{1}{3}$.

Another thing that captures our interest is that when the small firms' recruiting capacity k expands within the range $(\frac{1}{3}, \frac{1}{2}]$, the wages offered by the small firms increase while the wage offered by the big firm decreases. The intuition here is relatively easy to grasp. The growing capacities of the small firms intensify the wage competition between the two, which makes it more expensive for the big firm to recruit more workers. Here, an interesting result can be inferred from its reverse. That is, as the recruiting capacities in the small employers shrink, there is a wage increase at the big employer and a wage cut at the small employers.

However, the discussion would make more sense if we considered the interaction from workers. If the workers can form a reliable expectation of the net income (i.e. the wage posted by the firms net of the transportation cost that would be incurred), they would possibly choose to migrate from one city to another in the first place in order to enjoy a higher expected net income. Motivated by this, in the next subsection we introduce a three-stage game by allowing the workers to move in the first stage and then the big firm and the small firms sequentially set wages in the second and third stage.

4.2.2 A Three-stage Game

In this subsection, we consider a three-stage game between workers and oligopsonists in the similar context to the above wage-setting game. The setup of this three-stage game is formulated as follows:

In the three-stage game we take into account the possible migrating decisions of the workers. In order to facilitate the analysis, we only look at the case of two-point distribution, i.e. there are α workers initially residing in City A and the other $(1-\alpha)$ workers in City B. Here, all the workers are homogeneous and with the same reservation wage w_0. Additionally, we assume that *commuting* from one city to another requires a cost of c, while *permanent migration* from one city to another requires a one-off cost of s. As in the previous setup, the two co-locating firms in City A are constrained by recruiting capacity k but Firm B is not. The game is formulated in the following sequence[4].

(i) In the first stage, workers are allowed to relocate themselves at a one-off cost of s. Thus, a labour market is formed with β workers residing in City A and the other $(1-\beta)$ workers in City B.

(ii) In the second stage, observing the distribution of the workers in the market, the big firm decides its wage offer.

(iii) In the third stage, the two co-locating firms simultaneously and non-cooperatively set wages.

The key point here is that as the workers have perfect knowledge about the wage-posting strategy by the firm in the following stages, some of them would choose to relocate themselves in the first stage for higher (expected) payoffs. Here, we need to note that net wage includes: (1) the wages set by the firms; (2) the commuting cost that would be deducted from wage, if any; and (3) the migration cost incurred, if any. Next, we use backward induction by looking at the wage-setting game first and then coming back to the migration decision of the workers.

First, we have a short discussion of the case when $\beta \leq k$, i.e. the individual recruiting capacity of the co-locating firms exceeds the size of all local workers. It is relatively easy to comprehend that, when Firm B keeps all its local workers in equilibrium, the two co-locating firms in City A would always try to overbid its local rival to avoid nil labour supply. More specifically, when $(1-\beta)c \geq (1-2k)(1-w_0)$ Firm B pays $(1-c)$ for all its local workers and the two co-locating firms post a common wage that equals the marginal revenue $P = 1$ in equilibrium. When $(1-\beta)c < (1-2k)(1-w_0)$ Firm B chooses to pay w_0 for $(1-2k)$ local

[4]The sequence of the three-stage game is in line with the fact that it's much easier for the employers to adjust the wage offers than the employees' response of relocating.

workers; meanwhile, each firm in City A pays $(w_0 + c)$ for k workers. We need to note that $(2k - \beta)$ City B residents are daily commuters working for the firms in City A.

Secondly, we examine the case when $k < \beta < 2k$, i.e. the size of the local workers in City A exceeds the individual recruiting capacity but not the aggregate capacities of the two co-locating firms. If the wage offered by Firm i in City A is less than the *net* wage from Firm B or the reservation wage, it attracts nobody; if Firm i offers a wage that is less than its local rival's but still higher than the *net* wage from Firm B, it recruits the residual local workers; if Firm i offers a wage that is higher than its local rival's or the *net* wage from Firm B, it would employ workers to its capacity k. Thus, the labour supply for the firms in City A can be summarised as follows.

$$L_{Ai,3-stage}(w_{Ai}) = \begin{cases} 0 & \text{if} \quad w_{Ai} < \max(w_B - c, w_0) \\ \beta - k & \text{if} \quad \max(w_B - c, w_0) \leq w_{Ai} \leq \min(w_B + c, w_{Aj}) \\ k & \text{if} \quad w_{Ai} > \min(w_B + c, w_{Aj}) \end{cases} \tag{4.16}$$

The wage-setting strategy here is similar to the results we have derived in the previous subsection. When one of the co-locating firms overbids a little bit, it would benefit from the jump in labour supply. Meanwhile, it does not necessarily need to pay as high as the marginal revenue since there is some residual labour supply locally due to the recruiting limit of its local rival. Therefore, it implies a similar *mixed strategy* to be adopted by firm i in City A, which we will specify in the following paragraphs.

If a firm in City A picks the bottom wage $\underline{w}$, it will surely offer a lower wage than its counterpart. Also, since mixed strategy equilibrium requires that all the participants have *positive* expected payoffs, it implies that the net wage of Firm B for the residents in City A will not exceed the bottom wage offered by the firms in City A. That is

$$\max(w_B - c, w_0) \leq \underline{w}_A \tag{4.17}$$

It implies that the firm in City A recruits $(\beta - k)$ when it picks the bottom wage. Thus, the profit function can be written as

$$\pi_{A,3-stage}(\underline{w}_A) = (\beta - k)(1 - \underline{w}_A) \tag{4.18}$$

Profit maximisation implies

$$\underline{w}_A = \max(w_B - c, w_0) \tag{4.19}$$

It follows that

$$\pi_{A,3-stage}(\underline{w}_A) = (\beta - k)(1 - \max(w_B - c, w_0)) \tag{4.20}$$

Similarly, if the firm picks the top wage $\bar{w}_A$, it will surely offer a higher wage than its counterpart. The profit function can be written as

$$\pi(\bar{w}_A) = k(1 - \bar{w}_A) \tag{4.21}$$

One feature of mixed strategy equilibrium is that each player is indifferent among all the actions that he or she selects conditional on mixed strategy space of the counterparts. We therefore equate (4.20) with (4.21), and obtain the expression of $\bar{w}_A$, which is

$$\bar{w}_A = 1 - \frac{(\beta - k)(1 - \max(w_B - c, w_0))}{k} \tag{4.22}$$

So far, we have derived the lower and upper bound of the mixed strategy for the firms in City A. Next, our task is to find the distribution of the mixed strategy. Recall from (4.17) that the net wage of Firm B for the residents in City A will not exceed the bottom wage offered by the firms in City A. However, we have no idea whether the net wage offered by the firms in City A for the residents in City B exceeds the wage of Firm B or not. It needs to be discussed in the following two scenarios.

Scenario (i): $\bar{w}_A < w_B + c$.

By using (4.17), it follows that $w_B - c \leq \underline{w}_A \leq \bar{w}_A < w_B + c$. Thus, by using (4.16) the general form of the profit function for the firm in City A is

$$\pi_{A,3-stage}(w_A) = H(w_A)k(1 - w_A) + [1 - H(w_A)](\beta - k)(1 - w_A) \tag{4.23}$$

where $H(\cdot)$is the c.d.f. of the mixed strategy

Since each player has the same payoffs among all his/her actions in mixed strategy equilibrium, we can equate (4.23) to (4.20). Thus, given the Firm B's wage, the c.d.f. of the mixed strategy for firms in City A can be derived as

$$H(w_A) = \frac{(\beta - k)(w_A - \max(w_B - c, w_0))}{(2k - \beta)(1 - w_A)} \tag{4.24}$$

Scenario (ii): $\bar{w}_A \geq w_B + c$.

We note that, in this case, the firm in City A would grasp some of the workers from City B to its recruiting capacity when its net wage exceeds the wage of Firm B. By using (4.16), the profit function for the firm in City A can be written as

$$\begin{aligned}\pi_{A,3-stage}(w_A) &= (1 - H(w_A))H(w_B + c)(\beta - k)(1 - w_A) \\ &\quad + [1 - (1 - H(w_A))H(w_B + c)]k(1 - w_A)\end{aligned} \tag{4.25}$$

It seems a bit complicated, but we can manipulate it by letting w_A be at the lower bound $\underline{w}_A$, and equate it with (4.18). It yields

$$(\beta - k)(1 - \underline{w}_A)H(w_B + c) = (\beta - k)(1 - \underline{w}_A)$$

$$\Rightarrow H(w_B + c) = 1 \tag{4.26}$$

Recall our presumption here is that $\bar{w}_A \geq w_B + c$, thus, we can conclude that $\bar{w}_A = w_B + c$, which implies an atom of probability at the upper bound of the mixed strategy space. The intuition here is easy to understand: when a firm in City A picks the top wage it will surely recruit workers to its capacity, and there is no point of making higher wage offer to gain more workers from City B. Also, recalling from (4.17) that Firm B attracts no residents in City A in equilibrium, it is relatively simple to arrive at the following corollary.

Corollary 4.2 *For $k < \beta < 2k$, all the workers are employed by their local employer(s) in equilibrium.*

By using Corollary 4.2, the profit function for Firm B can be written as

$$\pi_{B,3-stage}(w_B) = (1-\beta)(1-w_B)$$

Since $w_B \geq w_0$, in equilibrium, Firm B will offer a wage that equals the reservation wage of all the workers.

Thus, the lower bound of the mixed strategy for the firms in City A also equals to w_0 by using (4.19). Now, the equilibrium wage-posting strategy for the firms in City A can be summarised in the following proposition.

Proposition 4.3 *i) If $\beta > \frac{k(2-2w_0-c)}{1-w_0}$, then the firms in City A randomise their wages within the interval $[w_0, 1-(\beta-k)(1-w_0)/k]$ following a c.d.f. of $H(w_A) = \frac{(\beta-k)(w_A-w_0)}{(2k-\beta)(1-w_A)}$;*

ii) If $\beta \leq \frac{k(2-2w_0-c)}{1-w_0}$, then the firms in City A randomise their wages within the interval $[w_0-c, w_0+c)$ following a c.d.f. of $H(w_A) = \frac{(\beta-k)(w_A-w_0)}{(2k-\beta)(1-w_A)}$; additionally, the distribution has an atom probability of $(1-H(w_0+c))$ at its upper bound of $\bar{w}_A = w_0+c$.

Proof. The mixed strategy for the firms in City is atomless in Scenario (i) when $\bar{w}_A < w_B + c$. By using (4.22), it follows that

$$1 - \frac{(\beta-k)(1-\max(w_B-c, w_0))}{k} < w_B + c$$

$$\Rightarrow 1 - \frac{(\beta-k)(1-w_0)}{k} < w_0 + c$$

$$\Rightarrow \beta > \frac{k(2-2w_0-c)}{1-w_0}$$

The mixed strategy for the firms in City A has an atom at its upper bound in Scenario (ii) when

$$\bar{w}_A \geq w_B + c$$

$$\Rightarrow \beta \leq \frac{k(2-2w_0-c)}{1-w_0}$$

Also, we note that (4.23) and (4.25) are exactly the same if we substitute (4.26) into (4.25). Then the c.d.f. within $[w_0, w_0+c)$ is identical to (4.24). It is relatively easy to derive that the atom probability of the upper bound (i.e. $\bar{w}_A = w_0 + c$) is $(1-H(w_0+c))$. ■

Lastly, we turn to the case when $\beta \geq 2k$. It is relatively easy to check that all three firms set a common wage that equals the reservation wage for the workers.

So far, the wage-posting strategy by firms in both City A and B has been well characterised. We introduce the following numerical example to illustrate the comparative statics with respect to the *realised* two-point distribution in the second stage.

Example 4.2 *Given the recruiting capacity $k = 0.4$ for each of the firm in City A, the commuting cost $c = 0.1$ and the reservation wage $w_0 = 0.6$, by observing the labour distribution (i.e. β in City A, $(1-\beta)$ in City B), the equilibrium wage(s) offered by the firms in both cities are displayed in Figure 4.7.*

When $\beta \leq 0.2$ all firms recruit the local workers. Firm B pays $(1-c)$ for $(1-\beta)$ workers and the two co-locating firms in City A post a common wage that equals the marginal revenue of product (i.e. $w_{A1} = w_{A2} = 1$).

When $0.2 < \beta \leq 0.4 = k$ Firm B pays w_0 for $(1 - 2k)$ local workers and each firm in City A pays $(w_0 + c)$ for k workers. $(0.8 - \beta)$ City B residents are daily commuters working for the firms in City A.

When $k = 0.4 < \beta < 0.8 = 2k$, there is residual labour supply for the lower-wage firm in City A, thus the two co-locating firms are in a Bertrand-Edgeworth situation and adopt mixed strategy and Firm B adopts pure strategy. It can be seen from Figure 4.7 that, given a specific distribution of the workers (i.e β in City A and $(1 - \beta)$ in City B), both firms in City A randomise their wage offers between the lower bound (i.e. the red solid line segment within $(0.4, 0.8)$) and the upper bound (i.e. the kinked dark line within $(0.4, 0.8)$). Additionally, the expected wage offered by the firm in City A is indicated by the dashed curve in Figure 4.7. The wage offered by Firm B is represented by the horizontal red line in Figure 4.7.

When $\beta \geq 0.8 = 2k$, all three firms offer a common wage that equals w_0.

We observe that the (expected) wages offered by all the firms are weakly declining in β, and we anticipate a similar result if we extend the two-point distribution. The intuition here is that as the recruiting capacities for the firms in City A are fixed, more workers residing in City A would release the wage competition between the two co-locating firms. Moreover, since the wage offers are strategic complements, the further consequence would be that Firm B would also lower its wage but still remain as attractive as before.

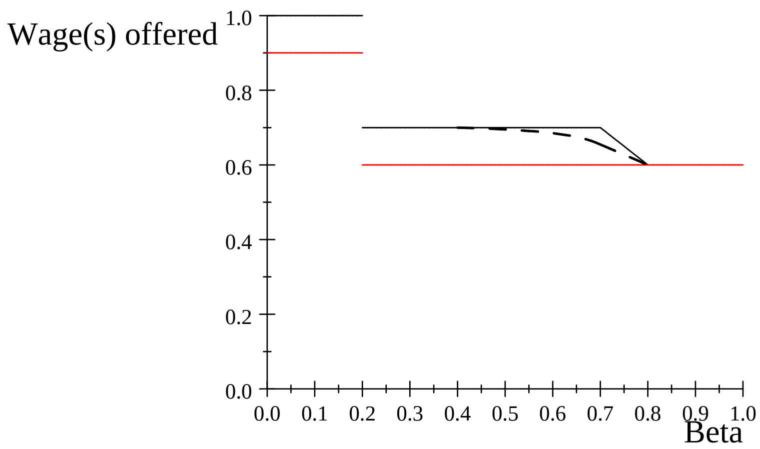

Figure 4.7: The wage(s) offered by firms in equilibrium when $k = 0.4$, $c = 0.1$, and $w_0 = 0.6$

So far in Proposition 4.3 we have derived a full characterisation of equilibrium strategies of wage posting for all the firms in the second and third stages. Next, we move on to the first stage to examine whether the workers have incentive to relocate themselves in a different city, given they have perfect knowledge about the wage-setting strategies for their prospective employers.

To begin with, we need to derive the (expected) wage **received** by workers.

Recall that when $\beta \leq k$ all the firms adopt pure strategies, then the (net) wage received by a worker equals the wage offered by his/her local recruiter(s).

When $k < \beta < 2k$ the firms in City A randomise their wage offers, there is uncertainty in the expected wages for the workers in City A. Recall from Corollary 4.2 that all the workers are to be recruited by the local employer(s). On the one hand, the workers in City B receive w_0. On the other hand, every worker in City A has a probability of $\frac{k}{\beta}$ of being offered a higher wage which follows a c.d.f. of $H^2(w_A)$ (i.e. the second order statistic), and has a probability of

$(1-\frac{k}{\beta})$ to be paid a lower wage which follows a c.d.f. of $[1-(1-H(w_A))^2]$ (i.e. the first order statistic). Thus, the expected wage received by the residents in City A can be written as

$$Ew_A(\beta,k)=\begin{cases}\frac{k}{\beta}\int_{w_0}^{\bar{w}_A}tdH^2(t)+(1-\frac{k}{\beta})\int_{w_0}^{\bar{w}_A}td(2H(t)-H(t)^2) & \text{if} \quad \beta>\frac{k(2-2w_0-c)}{1-w_0}\\ \\ H^2(w_0+c)[\frac{k}{\beta}\int_{w_0}^{w_0+c}tdH^2(t)+(1-\frac{k}{\beta})\int_{w_0}^{w_0+c}td(2H(t)-H(t)^2)] & \\ +2H(w_0+c)(1-H(w_0+c))[\frac{k}{\beta}(w_0+c)+(1-\frac{k}{\beta})\int_{w_0}^{w_0+c}tdH(t)] & \text{if} \quad \beta\leq\frac{k(2-2w_0-c)}{1-w_0}\\ +(1-H(w_0+c))^2(w_0+c) & \end{cases}$$

where $\bar{w}_A=1-\frac{(\beta-k)(1-w_0)}{k}$ and $H(t)=\frac{(\beta-k)(t-w_0)}{(2k-\beta)(1-t)}$

The (expected) wage received by the workers in City A is represented by the dotted curve in Figure 4.8. We need to note that, in the mixed strategy context, the expected wage *received* (i.e. the portion of dotted curve within $(k,2k)$) is different from the expected wage *offered* (i.e. the dashed curve) in Figure 4.7.

When $\beta\geq 2k$ the firms in City A would set a common wage that is attractive to the local workers in the third stage. Having full knowledge of this, Firm B would set a wage that equals the reservation wage in second stage. Thus, all the firms posts a common wage of w_0 in equilibrium, which is also the wage received by all the workers.

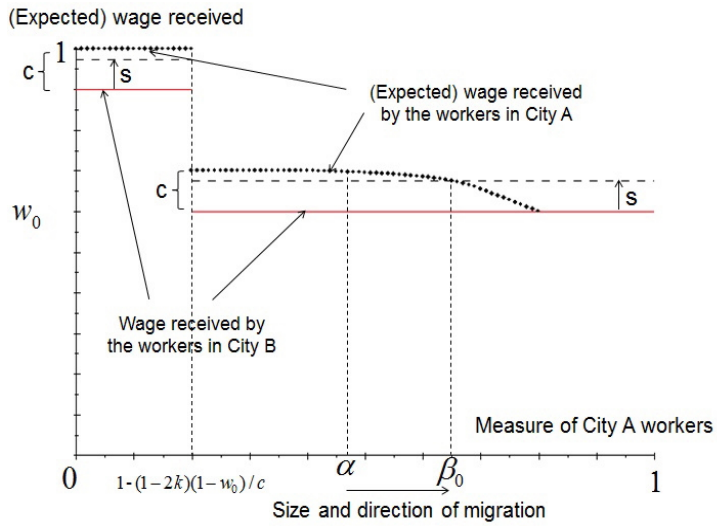

Figure 4.8: The (expected) wage received by the workers in City A and City B

Next, we look for the condition under which there is migration.

In Figure 4.8, we shift upward the red horizontal lines which represent the wage received by the workers in City B by s, the one-off permanent migrating cost. When $s<c$, the dotted curve and the dashed line intersect at β_0, where β_0 is implicitly defined by:

$$Ew_A(\beta_0,k)=w_0+s$$

Now we are able to arrive at the following proposition.

Proposition 4.4 *[i] If there is migration, it is from City B to City A;*

[ii] $s<c$ and $\alpha<\beta_0$ are necessary conditions for migration.

Proof. It is easy to comprehend that when $s \geq c$ no worker has the motivation to migrate in the first place, since he/she can always enjoy a higher net wage by paying the daily commuting fees rather than paying a higher one-off migrating cost.

Also, when $\alpha \geq \beta$ the expected gain by migrating turns out to be too costly to recoup the migrating cost, thus, all the workers choose to stay. ■

Furthermore, it can also readily be implied from Figure 4.8 that there are $(\beta_0 - \alpha)$ workers migrating from City B to City A when $s < c$ and $\alpha < \beta_0$. And it coincides with our intuition that increase in the daily commuting cost (c) and/or decrease in the permanent migrating cost (s) encourage more migration, if any.

4.2.3 Implication of Wage Dispersion and Trend of Wage Disparity

We have introduced two simple frameworks of the oligopsonistic competition in the above discussion. However, the value of the oligopsony approach to the labour market should ultimately be judged by its soundness in understanding and explaining the major issues in the labour market. Next, we put forward some areas where our models of oligopsonistic competition have helped to improve understanding.

"Law of one wage", which means that there is a single market wage for a given number of workers, is the key prediction of perfectly competitive labour market. However, voluminous empirical literature finds evidence of substantial wage dispersion among workers doing the same job in the same city (but not necessarily for the same employer). In the search theoretical context, the wage dispersion can be attributed to the search friction as well as the "competence" parameter (Postel-Vinay and Robin, 2002), i.e. heterogeneous worker abilities and firm productivities. In the two-firm duoposony framework, symmetry in firm productivity yields identical wages for all workers, and wage dispersion only rises in the case of the asymmetric productivity case (Bhaskar, Manning, and To, 2002). Now we might wonder whether the diversity in firm productivity is the only reason for wage dispersion, but apart from that, we do observe wage dispersion among homogeneous workers in the firms of similar productivity in the real world.

The wage dispersion issue for the firms with identical productivity can be handled conveniently with the setup of the wage-setting game in this paper. Recalling the setup in our oligopsony model that the small firms are identical but constrained by recruiting capacities, these employers do not necessary engage themselves in cutthroat wage competition since they may enjoy the residual labour supply which cannot be met by its rival. On the other hand, it's always beneficial to overbid, thus the small firms would end up adopting a mixed strategy in equilibrium. As a matter of fact, the wage dispersion in our model rises from the randomisation in wage-posting strategy by the small firms. The assumption of recruiting capacity constraints is in line with the reality of many private firms in the marketplace.

In spite of the wage dispersion amongst the small identical employers, we also address the trend of wage disparity between small and large employers as the recruiting capacities vary. We show that as the recruiting capacity shrinks there is a wage cut instigated by the small employers, whilst there is a pay rise in big employers. This is where there is scant theoretical or empirical analysis, but where there is interest since it is in line with the trend of wage disparity, especially as a consequence of economic downturns. The intuition here is that the reduced recruiting capacity mitigates the wage competition and drives down the wages of small employers, and in this way it reduces the marginal expenditure of labour of the competitor with no recruiting limit. As a result, the large firm absorbs more workers and meanwhile its wage goes up. However, we do not want to over-exploit the contribution of this result and hence only regard this as the partial explanation for the trend of wage disparity, since there might be other underlying factors such as implicit benefit received in large firms.

4.3 Conclusion

We have shown that in a spatial market of oligopsonists, two co-locating small firms with recruiting capacity constraints and a large firm without such limit are all competing for workers with horizontal differentiation. We have considered two frameworks.

In the wage-setting game, we use a uniform distribution of the workers along the "strip" market and elaborate upon the equilibrium wage-posting strategies adopted by all the firms, taking the recruiting capacity of the small firms as the only parameter. We show that when the recruiting limit is below some critical level (i.e. $k \leq \frac{1}{3}$) there are pure strategies of wage-posting for both the small and large firms (i.e. the small firms post a common monopolistic wage of $2k$, as though they can combine their capacities together, while the big employer sets wage at $\min(1-2k, \frac{1}{2})$). When the recruiting limit exceeds some critical level (i.e. $k > \frac{1}{3}$), there is a sudden jump in the strategies of all firms at $k = \frac{1}{3} + \varepsilon$, where ε is small and positive. As the recruiting capacities expand, the big firm cuts its wage while the small ones raise theirs. The intuition here is that the growing capacities of the small firms intensify the wage competition among themselves, which makes it more expensive for the big firm at the far end to recruit more workers through negative externalities. Besides, the reverse of this conclusion sheds some light on the recent trend of wage disparity between the private and public sector that has occurred during economic recession.

While it is relatively easy to understand the wage-setting model and characterise the equilibrium in the context of a spatial market, it is important to note that the physical distribution of the agents is actually a metaphor of heterogeneous preferences over the jobs, or the heterogeneous abilities to fulfill a job. Accordingly, the traveling cost can be interpreted as either a physical or psychological cost that needs to be overcome, and the permanent migration cost can be interpreted as a one-off training fee (borne by workers) to top up the qualifications needed for the jobs.

There are both merits and deficits in selecting the wage-setting game or the three-stage game. As we mentioned, the former can be applied in a more general context (i.e. the Hotelling preferences), while the latter is more specific as we interpret the Hotelling line as the real traveling space. And the setup of the three-stage game enables us to consider the possible interaction of the other side of the market (i.e. the decision of permanent migration for workers). Also, it is true that a prior bimodal distribution of workers can also be applied in the wage-setting model and would make more sense than the uniform distribution we have used. However, this simplified assumption has helped us to have some idea about which wage-posting strategies are employed by all the firms in the market. In the three-stage game, for the sake of brevity, we have assumed a two-point distribution of homogeneous workers, which helps us to concentrate on the interaction of the firms' wage-setting strategies and the workers' migration decisions. Besides, in some sense, the two-point distribution assumption does represent an extreme case of the bimodal distribution and as such merits investigation.

For brevity's sake, we have not explicitly explained the case of asymmetric assignment of recruiting capacities nor the competition among more than three firms. In our viewpoint, the solution of the three-oligopsonist case is already enough to shed some light on the wage competition among employers. With regard to the unequal capacities, the solution appears to be more of complex mathematical question than of particular economic interest.

Other interesting points include the discussion of the coordination issue amongst workers' migration - identification of the most probable equilibrium that would rise. Furthermore, another interesting question is to look at how much of the intuition captured in this paper can be carried over to a more general market, where the workers are bimodal distributed in the horizontal differential space, and how the wage bidding strategy of the firms will induce permanent reallocation of the workers. Our model is clearly a starting point for such further analysis.

Bibliography

[1] Atakan, A. E. 2006. "Assortative matching with explicit search costs". *Econometrica* 74(3): 667–680.

[2] Anderson, S. P., A. de Palma, and J. F. Thisse. 1992. *Discrete Choice Theory of Product Differentiation.* The MIT Press, Cambridge, MA

[3] Andreoni, J. 1990. "Impure altruism and donations to public goods: a theory of warm-glow giving". *The Economic Journal* 100(401): 464–477.

[4] Armstrong, M. 2006. "Competition in Two-Sided Markets". *The RAND Journal of Economics* 37(3): 668-691.

[5] Becker, G. 1973. "A Theory of Marriage, Part I." *Journal of Political Economy LXXI*, 813-847.

[6] Becker, G. 1974. "A Theory of Marriage, Part II." *Journal of Political Economy LXXII*, S11-S26.

[7] Beckmann, M. 1965. "Edgeworth-Bertrand Duopoly Revisited", *Operations Research* Vol. III: 55-68.

[8] Bertrand, J. 1883. "Review of Cournot's 'Rechercher sur la theorie mathematique de la richesse'", *Journal des Savants*: 499-508.

[9] Besanko, D., M. K Perry, and R. H Spady. 1990. "The logit model of monopolistic competition: Brand diversity". *The Journal of Industrial Economics* 38(4): 397–415.

[10] Bhaskar, V., A. Manning, and T. To. 2002. "Oligopsony and monopsonistic competition in labor markets". *The Journal of Economic Perspectives* 16(2): 155–174.

[11] Burdett, K., and M. G Coles. 1997. "Marriage and Class". *Quarterly Journal of Economics* 112(1): 141–168.

[12] Burdett, K., and D. T Mortensen. 1998. "Wage differentials, employer size, and unemployment". *International Economic Review*: 257–273.

[13] Chade, H. 2001. “Two-sided search and perfect segregation with fixed search costs”. *Mathematical Social Sciences* 42(1): 31–51.

[14] Dasgupta, P and Maskin, E. 1986. "The Existence of Equilibrium in Discontinuous Economic Games, I: Theory", *The Review of Economic Studies*, 53(1): 1-26

[15] Dixon, H. 1984. "Existence of Mixed Strategy Equilibria in a Price-Setting Oligopoly with Convex Costs", *Economic Letters* 16: 205-212.

[16] Dolan, R. J. 1995 "How Do You Know When the Price is Right?" *Harvard Business Review* 73: 174-183.

[17] Giraud-Heraud, E., H. Hammoudi, and M. Mokrane. 2003. “Multiproduct firm behaviour in a differentiated market”. *Canadian Journal of Economics* 36(1): 41–61.

[18] Hagiu, A. 2007. “Merchant or two-sided platform?” *Review of Network Economics* 6(2): 3.

[19] Hirata, D. 2009. “Asymmetric Bertrand-Edgeworth oligopoly and mergers”. *The BE Journal of Theoretical Economics* 9(1): 22.

[20] Hotelling, H. 1929. “Stability in competition”. *The economic journal* 39(153): 41–57.

[21] Jain, D. C, N. J. Vilcassim, and P. K. Chintagunta. 1994. "A Random-Coefficients Logit Brand-Choice Model Applied to Panel Data", *Journal of Business & Economic Statistics*, 12, issue 3, p. 317-328

[22] Kreps, D. M, and J. A Scheinkman. 1983. “Quantity precommitment and Bertrand competition yield Cournot outcomes”. *The Bell Journal of Economics*: 326–337.

[23] Lall, S and H. Selod, Z. Shalizi. 2006. World Bank Policy Research Working Paper No. 3915

[24] Levitan, R., and M. Shubik. 1972. “Price duopoly and capacity constraints”. *International Economic Review* 13(1): 111–122.

[25] Lu, X., and R. McAfee. 1996. “Matching and expectations in a market with heterogeneous agents”. *Advances in Applied Microeconomics*, vol. 6: 121-156

[26] Maskin, E. 1986. “The existence of equilibrium with price-setting firms”. *The American Economic Review* 76(2): 382–386.

[27] Nocke, V., K. Stahl, and M. Peitz. 2007. “Platform ownership”. *Journal of the European Economic Association* 5(6): 1130–1160.

[28] Oster, E. 2005. “Hepatitis B and the Case of the Missing Women”. *Journal of Political Economy* 113(6): 1163-1216

[29] Parker, G. G, and M. W Van Alstyne. 2005. “Two-sided network effects: A theory of information product design”. *Management Science*: 1494–1504.

[30] Pissarides, C. 1990. *A. Equilibrium Unemployment Theory*. Blackwell, Oxford

[31] Postel–Vinay, F., and J. M Robin. 2002. “Equilibrium wage dispersion with worker and employer heterogeneity”. *Econometrica* 70(6): 2295–2350.

[32] Rochet, J. C, and J. Tirole. 2003. “Platform competition in two-sided markets”. *Journal of the European Economic Association* 1(4): 990–1029.

[33] Rochet, J. C, and J. Tirole. 2006. “Two-sided markets: a progress report”. *The RAND Journal of Economics* 37(3): 645–667.

[34] Roth, A., and M. Sotomayor. 1990. *Two Sided matching: A Study in Game-Theoretic Modelling and Analysis.* Cambridge University Press, Cambridge, UK.

[35] Rubinstein, Ariel, and Asher Wolinsky. 1987. “Middlemen”. *The Quarterly Journal of Economics* 102(3): 581-594.

[36] Rust, J, R. Hall. 2003. "Middlemen versus Market Makers: A Theory of Competitive Exchange". *Journal of Political Economy*, 111, 353-403.

[37] Salop, Steven C. 1979. "Monopolistic Competition with Outside Goods." *The Bell Journal of Economics*: 141-156.

[38] Sattinger, M. 1995. “Search and the efficient assignment of workers to jobs”. *International Economic Review*: 283–302.

[39] Shimer, R., and L. Smith. 2000. “Assortative matching and search”. *Econometrica* 68(2): 343–369.

[40] Smith, S.P. 1976. "Pay Differentials Between Federal Government and Private Sector Workers", *Industrial and Labor Relations Review* 29: 179–197

[41] Smith, L. 2006. "The Marriage Model With Search Frictions". *Journal of Political Economy*, Vol. 114: 1124-1144.

[42] Smith, H., and D. Hay. 2005. “Streets, malls, and supermarkets”. *Journal of Economics & Management Strategy* 14(1): 29–59.

[43] Spulber, Daniel F. 1996. “Market Making by Price-Setting Firms”. *The Review of Economic Studies* 63(4): 559-580.