Inauguraldissertation

zur Erlangung des Grades eines Doktors der

Wirtschaftswissenschaften

Dissertation on Competitive and Directed Search

vorgelegt von:

Sheng BI

an der Fakultät für Wirtschaftswissenschaften

der Universität Bielefeld

October 2015

Declaration

This thesis has been written within the European Doctorate in Economics - Erasmus Mundus (EDEEM), with the purpose to obtain a joint doctorate degree in economics at the Department of Economics of Université Paris 1-Panthéon Sorbonne and at the Department of Business Administration and Economics at Universität Bielefeld.

Sheng BI

October, 2015

Printed on non-ageing paper DIN-ISO 9706

PhD Committee

Advisors:

Prof. François LANGOT, Université Le Mans

Prof. Dr. Bernhard ECKWERT, Universität Bielefeld

Other members:

Prof. Etienne LEHMANN, Université Paris 2 Panthéon-Assas , Referee

J.-Prof. Dr. Anna ZAHARIEVA, Universität Bielefeld

Prof. Antoine D'AUTUME, Université Paris 1 Panthéon-Sorbonne

Prof. Olivier L'HARIDON, Université de Rennes 1, Referee

To my loving grandparents and my parents

Acknowledgements

This PhD thesis was written during my studies in Université Paris 1 Panthéon-Sorbonne and Universität Bielefeld. I am grateful to European Doctorate in Economics Erasmus Mundus (EDEEM), funded by the European Commission's Education, Audiovisual and Culture Executive Agency. I am also very thankful for the generous grants from Université Paris 1 and the TIQE program under Deutsch-Französische Hochschule.

A number of people have contributed greatly to the making of this dissertation. First of all, I would like to thank my thesis advisors. I am very grateful to Prof. François Langot. Thank you for your pedagogical guidance and patience during these years when I was in Paris 1 and in Bielefeld. Thank you for all the helps and concerns whenever I have difficulties, and especially for the precious time and support that you have granted on me. I am also very grateful to Prof. Dr. Bernhard Eckwert. Thank you for kindly accepting me under cotutelle, giving me important suggestions and detailed comments during discussions, and providing me super environment for me to focus on my research. I also owe my special thanks to Prof. Bertrand Wigniolle who always pays keen attentions to my situations, gives me so many supports and encouragements, provides me so many useful information, and I could never thank you too much.

In addition, I would like to express here, all my gratitude towards the other members of my PhD. Committee: Prof. Etienne Lehmann, Prof. Antoine D'Autume, J-Prof. Anna Zaharieva and Prof. Olivier l'Haridon, who took substantial time to review my work, provide valuable comments and suggestions. I know Prof. Lehmann and Prof. D'Autume since I was a master student in Paris 1, and am quite influenced by their ways of thinking and rigorous manner of treating problems. J-Prof. Zaharieva gives me a lot of important comments, is always available for me to talk with. I am acquainted with Prof. l'Haridon just before my pre-defense, but has already benefited greatly from his insights when he helps me improve the dissertation.

I take here the unique opportunity to express my gratitude and admiration to all my teachers and professors during my studies in University of Paris 1, University of Bielefeld and University of Autnoma Barcelona: Prof. François Langot, Prof. Dr. Bernhard Eckwert,, Prof. Bertrand Wigniolle, Prof. Etienne Lehmann, Prof. Antoine D'Autume, J.-Prof. Anna

Zaharieva, Prof. Jean-Olivier Hairault, Prof. Caballé, Jordi, Prof. Cabeza, Maite, Prof. Obiols, Francesc, Prof. Vilà, Xavier, Prof. Christoph Kuzmics, Prof. Herbert Dawid, Prof. Frank Riedel, Prof. Jean Imbs, and Prof. Nicolas Dromel. I also profited from numerous remarks during my participation to the T2M conference in Lyon, QED workshop in Vienna, GAINs seminar in Université le Mans, in the Chinese Academy of Agriculture organized by Prof. Lv Kaiyu. Macro-retreat in Paris, the annual meetings EDEEM Summer Meeting Jamboree 2014/2015 (Bielefeld/Venice), EBIM workshop, Macro workshop, as well as Bigsem colloquium.

I am very grateful to all EDEEM staff, especially Prof. Bertrand Wigniolle, J.-Prof. Anna Zaharieva, Dr. Ulrike Haake, and Prof. Herbert Dawid. They are always ready to help me and give me support during my stay in Paris and in Bielefeld. In Paris 1, I would like to thank also Natalie Louni, Elda André, Loïc SOREL, Liliana de Freitas, and Viviane Makougni, in Bielefeld, I would like to thank Sabine Kuhlmann, Helga Radtke, Christine Charalambous, Diana Grieswald-Schulz, and in Universitat Autonoma de Barcelona, I would like to thank Mercè Vicente, Àngels López, as well as Beviá Carmen for their kindness and their efficiency in providing helps in many administrative matters.

I have enjoyed a warm and stimulating international environment during my studies in Paris, Bielefeld and Barcelona, thanks to my colleagues and my friends: Lalaina, Shuai, Yu, Ni, Nan, Antoine, Lorenzo, Hongli, Peter, Nikolai, Andreas, Carla, Xin, Abhishek, Yuanyuan, Xinmian, Lan, Vincent, Camille, Axelle, Baris, Hamze, Alex, Liang, Kaiyu, Bertrand. I am very lucky to have you as friends, when I am having difficulties you are always around to help me. Thanks to all of you, my stay abroad was a very pleasant and enriching adventure where I have learnt a lot.

Last but not least, I would like to thank my family. I owe all my gratitude to my Grandpa, Grandma, my father, and my mother. You provide me firm supports so that I have this precious opportunity to study abroad, cultivate my philosophy, travel and discover so many wonderful places, meet so many good people, and experience diverse cultures, there are laughs and tears, which I could share with you. Many thanks to my aunts and uncles who treat me as if I am their own child. Thank you for all of your love, and encouragement. I am always so grateful to my grand-parents; they passed away during my study abroad, and I was not able to fulfill the responsibility to take care of you, and return at least partially the love that you have granted me for all the time. It is for me a lifelong regret. My parents and aunts are also so supportive with incredible understandings. I owe you so much, and thank you all for your constant attention and your prayer. I wish you all the best joys.

Sheng BI, Bielefeld, October 2015.

General Introduction

Searching for a trading partner is time-consuming, and can often end up unfruitful. In economic terms, the market is said to feature search and matching frictions. Prominent examples of contexts where such friction is quantitatively important are labor, housing, and marriage markets etc. Admitting such a friction, in search theory, we are concerned with to which extent can the resources be allocated efficiently.

From a macroeconomic perspective, Mortensen and Pissarides (1994) are the first to formalize this friction in an equilibrium model to study the labor market coexistence of unemployment and unfilled vacancies. In this traditional framework, the mechanism of determination of trading price and the mechanism how search friction takes place are independently framed. A direct consequence is the appearance of search externality, which drives the economy away from efficient allocation of labor.

The underlying reason of such inefficiency is the negligence of the important relationship between price and search behavior: observing the prices, an agent who chooses which price to visit is able to anticipate how intensively the others are going to frequent each price in question, and this anticipation allows her to calculate her own probability of successful trading. Moen (1997) finds that when agents are price-takers and maximize income subject to a given set of market parameters, the search externality can be internalized and efficiency can result. He provides informally a justification on how competitive search equilibrium can be implemented by the above relationship - a process with profit-maximizing firms' price posting, and workers' subsequent application to the observed prices that give them the highest expected utility from search.

Burdette, Shi and Wright (2001) take a first step in formalizing how the posted price can direct subsequent search. Their contribution consists in pointing out the distinction of a market where the number of agents is finite and the market where this number tends to be large, and deriving endogenously the matching function. A recent work from Manolis and Kircher (2012) provide preliminarily game theoretical foundation of these price posting models with subsequent search, and shows formally how competitive search equilibrium results when the number of agents tends to be large. Apart from the advancement of pure theory in this direction, there are few works which apply this theory to understand stylized facts, while this framework has a clear advantage in abstracting from the inefficiency arising from search externality, which allows us to focus solely on other market imperfections that are of our concern. Hence, in this dissertation, we take this price posting approach to study three issues in labor market.

The first issue concerns the premature quitting of workers in an employment relationship. Our framework is especially suitable for contexts such as disability shock, retirement, maternity leaves etc. Such premature quitting creates turnover risks for firms, hence the firms would like to propose wage profiles to minimize or avoid it. In this issue, the asymmetric information plays an important role. We adopt an approach of mechanism design and consider different timings when the private information is realized. In a follow-up paper, we propose a specific age-directed policy by which this inefficiency can be alleviated, and study its implication on welfare and aggregate output. In the second issue, we revisit classical welfare analysis of impact of discrimination on skill choice in the presence of search friction. Our main contribution is to consider a multi-dimensional hiring norm along both productivity-related and -unrelated characteristics. We show how strategic skill investment between favored and discriminated group endogenously arises. We compare also two wage determination mechanisms (posted and bargained wage) to check the robustness of our result. In the third issue, we consider to which extent can unemployment benefit and minimum wage correct inefficient allocations arising from firms' market power. Our context concerns small markets where the workers/firms ratio is not large. The significance of such small sub-markets is highlighted by Shimer (2007) who finds that the market can be small down to a tightness of on average 10 unemployed workers per firm when taking the "occupation, geographical area" as the sub-market characteristic. The market imperfection comes from the fact that in such a small labor market firms pay less than competitive level of wages. We proceed from an Industrial organizational perspective, and suggest focusing on both misallocation of labor and misallocation of surplus when analyzing the effectiveness of the policy instrument.

Table of contents

Chapter 1

Dynamic Contracts in Search Equilibrium under Asymmetric Information

Coauthored with Professor François Langot

Abstract

We consider a multi-period competitive search equilibrium where inefficiency is due to asymmetrical information. At the time when firms commit to specific hiring costs, workers hold private information on their intention of quitting before the termination of the contract. A long-term employment relationship leads to both higher wages and employment rate. The optimal separating contract is an ascending wage profile, which distorts the allocation of the workers who will quit later (the 'good' workers) in order to prevent the workers who will quit earlier (the 'bad' workers) from applying for these jobs. We also endogenize the separation decision (the extensive margin). Two cases are considered, where heterogeneity either happens ex ante or ex post stage. Apart from the similar ascending wage profile as in the benchmark, we establish that there is excessive separations compared to the full-information economy.

Keywords: Competitive search equilibrium, asymmetric information, dynamic contract, extensive margin

JEL classification: D82, D86, J14, J26, J64

1.1 Introduction

> The basic message of *[the Phelps volume published in 1970]* was that one could and should consider the *[labor market]* problems as the outcomes of market behavior of individual agents who act in their own interests as best they can in a market environment characterized by uncertainties and incomplete information.
> Mortensen (2011)

Since Phelps (1970), the search and matching frictions are presented as typical departures from the Walrasian theory of the labor market, allowing to account for the trading externalities. But the competitive search equilibrium leads to reducing the distance between the matching model and the Walrasian paradigm: a decentralized market can internalize the search externality, leading to an efficient equilibrium (Moen (1997)).[1] As in Guerrieri, Shimer and Wright (2010), (GSW), this paper re-introduces informational incompleteness in a competitive search environment to account for the inefficiency of the labor market allocation. The 'good' workers generate higher surplus, however, for employers who target such workers, they expect to wait longer to get matched with them, which implies higher virtual cost. Given their commitment on hiring strategy, these employers want to prevent undesirable ('bad') workers from applying, and it is achieved by sacrificing part of surplus with these 'good' workers. Then even if the discrimination process[2] is directed 'against' the bad, its induced cost (information rent) is 'paid' by the good - those who have incentive to work longer in our context.

The aim of this paper is to extend the GSW's results to a dynamic context. The intertemporal choices allow us to account for *(i)* the large turnover risks that are faced by the firms, and *(ii)* the risk aversion that characterizes the worker utility function leading them to ask for a smoothing of the wage profile. Thus, in our benchmark, the unobservable characteristic is workers' quitting probability, whereas in the static context of GSW, this is the labor productivity. Moreover, we analyze two types of asymmetrical information: *(i)* a classical ex ante heterogeneity where the employers' problem is to screen the undesirable types and to induce self-selection among the pool of ex-ante informed workers (adverse

[1] In competitive search equilibrium, the information on these ex-post matching returns is perfectly observable. Thus, an optimal segmentation occurs at the equilibrium: the 'good' and 'bad' workers are attracted to specific markets. This leads to the conventional view that the market gives optimally more chances to 'good' workers: the segmentation process is applied to the less profitable workers. One may then have the tendency to believe that the cost associated with these risks of heterogenous returns are borne by workers, the 'optimal discrimination' being borne by these 'bad' agents.

[2] We say the 'bad' workers are discriminated against in the sense that their expected utility from search is less than proportional compared to the good workers. We do not mean that the firms always prefer hiring bad workers to good workers. Firms are actually ex-ante identical, and indifferent to posting long term or short term contracts.

selection), *(ii)* an ex post heterogeneity where the worker observes a private information only after the employment relationship has been established, leading employer to propose contracts allowing her both to sort and retain workers (a commitment problem). We also extend the results from Guerrieri (2008) by taking into account of risk aversion.

From a methodological point of view, we describe our main results in a two-period stylized model of the labor market and also extend in the last section our benchmark to three periods to study wage growth. This extension allows us to account for the transitory choice of worker mobilities: after a separation, they can re-enter into the labor market. Fundamentally, there are two types of workers in the labor market: the ones that have a better outside option after a tenure of one 'period', and the others having this option only after a tenure of two 'periods'.[3] Given this heterogeneity, the search frictions can lead the employers to sort workers, since the ex-post returns for the firms are heterogeneous, even if the costs of posting a vacancy are homogenous. In the context of the ex ante heterogeneity, we show that GSW's results can be extended to this dynamic context. More precisely, the search and matching frictions ensure that the separating equilibrium can always be achieved. Indeed, with respect to a competitive environment *à la* Rothschild and Stiglitz (1976), (RS), the competitive search equilibrium introduces a new restriction: the matching is bilateral – each principal can serve at most one agent ex post, whereas in RS, each principal can serve the total population.[4] When the numbers of principals and agents are sufficiently large, any pooling deviation is proved to be unprofitable. Thus, the meeting probabilities adjust endogenously to clear the market. Hence GSW's study provides theoretical benchmark for our problem, and there are mainly two aspects along which our paper is different from theirs.

1. Firstly, our paper captures wage dynamics in a tractable way. In our model, asymmetric information not only affects the trade-off between ex post surplus and the trading probability, it also leads to distortion on the intertemporal wage dynamics originated from workers' risk aversion. As is well known in terms of contract theory, we show that workers are required to concede certain amounts of wages in the firm as guarantee at the beginning of the employment relationship, and in case they indeed stay with the

[3] If we interpret these 'periods' as the duration of the employment contract in a firm submitted to large turnover costs, the outside option is related to the workers' private preferences for the type of work she likes to do, location, alternative job opportunities etc... We model these diverse options as a reduced-form outside option. Alternatively, one can interpret this outside option as the realization of risks linked to maternity or to health, which is thus exogenous. Finally, if we think at the labor market of the older workers, one can interpret the outside option as the retirement value, considered as exogenous.

[4] This trading restriction acts as a sort of capacity constraint that plays the crucial role in resolving the RS non existence problem: when agents are not able to obtain a contract with probability 1 (due to coordination friction), they engage themselves actively in a subgame to maximize their expected payoff related to optimal search, which is the solution of a trade-off between the wage and the probability of obtaining it.

firm in the following period, the amount can be recouped along with the interest rate (tenure-increasing wage profile): workers should be responsible for the risk caused by themselves. The crucial question is then whether the firms are over- or under-compensating the workers under distortion caused by workers' private information. We prove that it is the former situation: under competitive search, risk neutral firms cut 1st-period wage to such an extent that bad types are disincentivized to apply. However, in order not to lose the applications from the good types, the 2nd-period wage has to increase to such an extent that the total lifetime wages given to good type workers manage to surpass the full-information level, i.e., firms over-compensate in terms of total wage transfer. Compared to the full-information case, firms have lower ex post surplus, but in exchange, they are to receive more often applications from workers due to the excessive wage amounts posted. Thus, our paper gives equilibrium foundations to the old result of Salop and Salop (1976), where an optimal compensation policy is such that employees are asked to make an up-front payment at the beginning of the employment relationship, which will be refunded gradually in the subsequent periods, conditional on their stay with the firm. Workers with a higher quit probability choose not to apply to firms with this contract offer, since they expect that they will have a low probability of receiving the repayments in the future. A self-selection mechanism is thus induced. It is also important to mention that the results from Salop and Salop are not robust enough: if workers are risk-averse, and if firms are risk-neutral on a competitive market with a free-entry condition, the Salop and Salop model immediately collapses to share the same problem as RS model: there may exist no equilibrium in such a competitive market, when the fraction of bad workers is sufficiently low in the economy. Within a competitive search equilibrium in the manner of GSW, this problem is naturally solved.

2. Secondly, our immediate extension features variations of the extensive margin. This is absent in GSW and most of the standard adverse selection literature. Risk averse workers differ in their outside option value, and choose whether to enter into a long-term or short-term employment relationship before search. Workers with lower outside option value wish to avoid non-employment risk, and in response, the market offers a high employment rate. However, this high job finding rate raises firms' vacancy risks. We then show that, under asymmetric information, workers choose to enter into short-term relationship more often because the introduce of incentive constraint discretely lowers workers' expected utility from entering into a long-term relationship. Also, the 'good' workers suffer from inferior employment rate compared to the case of full

information, which implies a higher average non-employment rate in the decentralized economy.

If we depart from the informational structure *a la* GSW, our stylized model of the labor market also extends the results of Guerrieri (2008) where workers only know their types after the acceptance of the contract. In this case, the contract should ensure that ex post the workers do not break the match unilaterally. Specifically, we are in a situation where firms have full commitment, whereas workers have limited commitment. Compared to Guerrieri (2008), our second extension features risk aversion and we focus on workers' outside option (denoted by h), which depends on permanent opportunities outside the firm, instead of matching specific disutility, conditional on a transitory specific match. We find that there is a markup on the term related to information rent, i.e. we have $\frac{G(h)}{g(h)}u^{'}(h)$ instead of $\frac{G(h)}{g(h)}$.[5] Standard literature assumes that $\frac{G(h)}{g(h)}$ is increasing, while $\frac{G(h)}{g(h)}u^{'}(h)$ clearly may not. Under risk aversion ($u^{''}(h) < 0$), the slope of the information rent is less steep. It suggests that risk averse workers are more conservative about the attractiveness of the outside option captured by $\frac{G(h)}{g(h)}$; in other words, the distortion of quitting threshold is less responsive to the degree of severity of the information problem (measured by the Lagrangian multiplier of the information constraint). However, it is clear that compared to the full-information economy, the quit rate is larger and the employment rate is lower.

The paper is organized as follows. In the first part of the paper (section 2), we present the benchmark where workers' quitting probability is known ex ante, exogenous, and not observable by the firms. There are two types of workers: those who quit sooner and the rest who quit later. As in GSW, we show that there always exists a separating contract, which is characterized by a distortion of the wage profile afforded by the good (those who quit later): in the segment of these workers, the posted wage profile is age-increasing as a direct consequence of the incentive compatibility constraint, whereas, In the segment of the bad, there is no distortion (flat wage rate). This suggests that the lack of information concerning the workers' horizon leads to sub-optimal welfare of the workers accepting a priori to work longer. The second part of the paper (section 3) shows how the dynamic wage contract induces additional distortion on the extensive margin of the labor market. Indeed, the decision on working horizon depends crucially on the outside option of the workers. Workers can be aware of this heterogeneity before (ex ante heterogeneity) or after (ex post heterogeneity) the matching process. We analyze these two types of heterogeneity. First, we show that the optimal contracts lead to introducing wage distortions (characterized by an age-increasing wage profile) on the labor market for the good workers, for both types of heterogeneity, as

[5] $G(h)$ is the CDF of the random variable h, $g(h)$ is the pdf of h, and $u^{'}$ is the marginal utility.

in the benchmark. In the case of ex ante heterogeneity (section 3.1), the new result is that the equilibrium reservation threshold is lower than in the full-information economy: the less transparent is the information on the outside options, the greater the number of workers that choose short-term employment relationship. With ex post heterogeneity (section 3.2), we also show that not only do more workers choose to quit the labor market early, but the employment rate of the workers who want to work for two 'periods' is lower than in the full information case. In the section 4, we discuss the robustness of our result in a multi-period extension of the benchmark. The section 5 concludes.

1.2 Competitive search with exogenous separations

Consider a 2-period economy with search friction. There is a fixed quantity of workers, as well as a pool of firms, the amount of which is to be determined in the equilibrium. Workers start unemployed, and can get hired by a firm through a job-search process. We assume that an employed may leave the firm before the termination of the employment relationship with certain probability, which is interpreted as probability of quitting[6]. There are two possible types of workers[7]: a fraction α remain in the second period with probability p_H, called type "H"; while the rest, with fraction of $(1-\alpha)$, are type "L" workers, who remain in the second period with probability p_L. We assume $p_H > p_L$.

All firms are ex ante identical. To start production, each has to create a vacancy with a sunk cost C, and then announce wages in hopes of getting matched with a worker.[8] When an employer hires a worker, the match produces y per period. We assume that firms and workers share the same subjective discount rate β.

There are matching frictions. We define q to be the tightness of the labor market, i.e. the workers/firms ratio.[9] The hiring probability for the firms is $H(q)$, while $F(q)$ denotes the workers' job finding rate. As is common in the matching models, we assume

- the constant return to scale relationship $H(q) = qF(q)$;
- $F(q)$ the workers' job finding rate is decreasing and concave in q, the expected number of workers per firm, i.e. $F'(q) < 0$ and $F''(q) < 0$;

[6]It would be useful to think of the following activities as examples:*(i)* mobilities toward another sector or toward another geographical area, *(ii)* non-participation due to disability or retirement

[7]Introducing more types does not alter the main results, we stick to this case for the ease of exposition and refer interested readers to GSW (2010) for discretely many types.

[8]Following Acemoglu and Shimer (1999), we make use of the concept of submarket, where each firm, although ex ante homogeneous, will choose which type of agents to attract.

[9]In terms of directed search, the queue length is the expected number of applicants from a job seeker's perspective, as in Acemoglu and Shimer (1999).

- $H(q)$ the firms' hiring probability is increasing and concave in q, i.e. $H'(q) > 0$ and $H''(q) < 0$;
- The elasticity of firms' hiring probability is a decreasing function in q and between $(0,1)$, i.e. $\varepsilon'(q) = \left[\frac{H'(q)}{H(q)/q}\right]' < 0$ and $\varepsilon(q) \in (0,1)$.

For each period, the wage posting game has three stages: (i) wage posting stage, where firms announce wages; (ii) job search stage, in which workers observe the posted wages, and choose where to apply; (iii) production stage, where the workers who get matched with firms produce, and payoffs are realized.[10]

1.2.1 The competitive search equilibrium under full information

In this section, we will describe the benchmark model, and show that with full information, workers who can build a long-term employment relationship (have higher probability of remaining in the firm) expect higher expected utility from search.

Workers. With full information, two types of workers search respectively in different submarkets opened to them. These two submarkets are indexed according to their types, i.e., H or L. At the beginning of the first period, each worker enters her submarket to search for a job. When a worker finds a job today, a type H (L) will remain at the second period with a probability of p_H (p_L); when she fails to find a job in the first period, she enjoys unemployment benefit z today, and could search again at the beginning of the second period. Since the second period is the end of the economy, any worker who searches for a job at second period works only for one period if a job is found, and otherwise she enjoys the unemployment benefit. Notice that if the worker leaves the market in the second period (with probability $(1-p)$), they receive utility from some outside option z_2. Since under such a formulation the value of z_2 does not alter our result, we normalize it to zero.[11] The worker's expected utility could be written in the following way where $t \in \{L,H\}$, $p \in \{p_L, p_H\}$, and

[10]Competitive search could be understood as directed search in a large economy, in the sense that when the number of agents tends to infinity, individual deviations are not able to alter the payoff expected from the market. Galenianos and Kircher (2012) demonstrated this result.

[11]In general, this uncertainty can be related to a 'negative' or a 'positive' shock. Examples for negative shock can be productivity drop, disability etc..., in which case $z_2 \leq z$. An example for the interpretation as a positive shock is job-to-job transition.

$1p$ stands for one-period:

$$\begin{aligned} U_t(p) &= F(q_t)\left[u(w_{t,1}) + \beta\left(pu(w_{t,2}) + (1-p)u(z_2)\right)\right] \\ &\quad + (1-F(q_t))\left[u(z) + \beta\left[pU_{1p} + (1-p)u(z_2)\right]\right] \end{aligned}$$

where

$$U_{1p} = F\left(q_{1p}\right)u(w) + \left(1 - F\left(q_{1p}\right)\right)u(z)$$

Firms. Ex ante identical firms choose which submarket to enter, and the corresponding wages. This implies that each firm can only attract one type of worker at most.[12] We assume that once the worker leaves the firm prematurely with probability $1-p$, the job is destroyed, the firm in question exits the market and is replaced immediately by another. The profit of firms is written as follows, where the first line is the profit for a firm contracting with a two-period worker, and the second line is the profit for a firm contracting with a one-period worker:

$$\begin{aligned} \pi_t(p) &= H(q_t)\left[(y - w_{t,1}) + \beta p(y - w_{t,2})\right] - C \\ \pi_{1p} &= H(q_{1p})(y - w_{1p}) - C \end{aligned}$$

Definition 1. *A competitive search equilibrium is a set of wage profiles $w^*(p) = \{w^*_{t,1}(p); w^*_{t,2}(p)\}$, together with a function $q^*(p) = q(w^*(p))$ and utility levels $U^*_t(p)$ for $p \in \{0, p_L, p_H\}$ that satisfy the following conditions:*

(1) employers' profit maximization and free entry,

$$H(q^*_t(p))\left[\left(y - w^*_{t,1}(p)\right) + \beta p\left(y - w^*_{t,2}(p)\right)\right] - C \leq 0$$

(2) workers' optimal job application,

$$\begin{aligned} U^*_t(p) &\geq F(q^*_t(p))\left[u\left(w^*_{t,1}(p)\right) + \beta\left(pu\left(w^*_{t,2}(p)\right) + (1-p)u(z_2)\right)\right] \\ &\quad + (1 - F(q^*_t(p)))\left[u(z) + \beta\left[pU_{1p} + (1-p)u(z_2)\right]\right] \end{aligned}$$

[12] If firms are allowed to attract both types of workers at the same time, firms are showing a portfolio problem where workers with higher value of p are always hired in preference to the workers with lower value of p. The problem is more complex, and we refer interested readers to Shi (2006) for more details.

where $U_t^*(p)$ is given by

$$U_t^*(p) = \max_{w(p)} \left\{ \begin{array}{l} F\left(q_t^*\left(w\left(p\right)\right)\right)\left[u\left(w_{t,1}\left(p\right)\right) + \beta\left(pu\left(w_{t,2}\left(p\right)\right) + \left(1-p\right)u\left(z_2\right)\right)\right] \\ + \left(1 - F\left(q_t^*\left(w\left(p\right)\right)\right)\right)\left[u\left(z\right) + \beta\left[pU_{1p} + \left(1-p\right)u\left(z_2\right)\right]\right] \end{array} \right\}$$

The solution is characterized by the following equations: for $p \in \{0, p_L, p_H\}$, we have

$$\begin{aligned} w_{t,1}^*(p) = w_{t,2}^*(p) &= w_t^*(p) \\ H\left(q_t^*(p)\right)\left(1+\beta p\right)\left(y - w_t^*(p)\right) &= C \\ \frac{1-\varepsilon\left(q_t^*(p)\right)}{\varepsilon\left(q_t^*(p)\right)} &= \frac{y - w_t^*(p)}{\frac{1}{u'\left(w_t^*(p)\right)}\left[u\left(w_t^*(p)\right) - \frac{1}{1+\beta p}\left(u\left(z\right) + \beta p U_{1p}^*(p)\right)\right]} \end{aligned}$$

where $\varepsilon(q)$ denotes the elasticity of the firms' hiring probability with respect to the queue length. Leaving aside the conditions of free entry, the last equality (from the first order condition on q) could be interpreted in the following manner: the LHS gives the relative contribution of firms and workers to the total number of matches created, as the ratio between the firms' search effort to the workers' search effort, whereas the RHS gives the surplus ratios between firms and workers, in units of goods (the surplus of the workers is thus corrected by the marginal utility).

By the comparison of allocations for the type-H and type-L workers, we have the following proposition, where we may use the superscript '$*$' and 'FI' interchangeably to stipulate that the allocation is achieved under full information (FI):

Proposition 1. *The competitive search equilibrium is characterized by:*

- *The optimal level of per period wage is such that $w_H^* > w_L^*$*
- *The optimal level of expected queue length is such that $q_L^* > q_H^*$*
- *The value from search is such that $U_H^* > U_L^*$.*
- *The wage increases less faster than the increment of productivity with respect to p: $(1+\beta p_H)w_H^* - (1+\beta p_L)w_L^* < (1+\beta p_H)y - (1+\beta p_L)y$.*

Proof. See Appendix 1.6.1. □

This proposition states that not only do the type H workers earn higher per-period wages, but they also get matched more often compared to the type L workers whose expected employment horizon is of shorter term ($p_L < p_H$).

There is one further observation. Consider a particular pair of parameter values by setting $p_H = 1$ and $p_L = 0$,[13] so that in the second period workers either remain in the firm for sure or leave for sure. Now suppose that a worker can choose to be involved in a long term or repeated short term relationship, then workers are better off involved in a long-term relationship. This result is formalized in the following proposition:

Proposition 2. *Take $p_H = 1$ and $p_L = 0$. Suppose that a worker can choose to be involved in a long-term (2 periods) or repeated short-term (1-period) relationship, then workers are better off involved in a long-term relationship in terms of value from search, i.e. $U_H^* > (1+\beta)U_L^*$.*

Proof. See Appendix 1.6.2. □

It is important to remark that although firms pay the same entry cost C on each sub-market, the effective cost, defined as $\frac{C}{H(q)}$, is larger than the unit cost C, because it is obtained by multiplying the average duration of the search. The inequality $q_L^* > q_H^*$ coincides with our intuition that in order to earn higher ex post surplus, the firms which post type H contracts anticipate that they have to wait longer till such a worker appears.

With perfect information, we have established that the workers who expect longer employment relationship (measured by p) receive higher expected utility from search. Although all different types offer the same quality of labor, type-L workers receive lower wages and suffer from a lower expected utility. It is this situation which generates incentives for the type-L workers to visit the submarket for the type-H workers, if information is not perfectly shared. Thus firms which post long-term employment contracts and are uninformed about the workers' types may suffer loss if they hire a type-L worker, despite the fact that they have incentives to hire a type-H worker. We are thus led to consider the possibility for the firms to manipulate wage profiles to induce self-selection among the workers, such that no workers will misrepresent their types.

1.2.2 The allocation under asymmetric information

In the previous subsection, we have shown that type-L workers have incentive to misrepresent their type, because they find that higher expected utility could be obtained if they send applications to a market where firms want to hire type-H workers. In order to discourage type-L workers from joining the submarket for type-H workers, an alternative wage profile should be posted to induce self-selection. A standard way is to introduce an incentive compatibility constraint, such that type-L workers find it unattractive to pretend to be a type

[13] In this case, the type L is equivalent to type $1p$

H. The optimization program is now written as follows where the index AI denotes the equilibrium allocation under asymmetric information:

$$
\begin{aligned}
U_H^{AI} &= \max_{\{q,w_{H,1},w_{H,2}\}} \left\{ \begin{array}{l} F(q_H)\left[u(w_{H,1})+\beta p_H u(w_{H,2})\right] \\ +(1-F(q_H))\left[u(z)+\beta p_H U_{1p}\right] \end{array} \right\} \\
\text{s.to} \quad & H(q)\left[y-w_{H,1}+\beta p_H(y-w_{H,2})\right]=C \\
\text{I.C.} \quad & \left\{ \begin{array}{l} (q_H)\left[u(w_{H,1})+\beta p_L u(w_{H,2})\right] \\ +(1-F(q_H))\left[u(z)+\beta p_L U_{1p}\right] \end{array} \right\} \leq U_L^{AI}
\end{aligned}
$$

where

$$
\begin{aligned}
U_L^{AI} &= \max_{\{q,w_{L,1},w_{L,2}\}} \left\{ \begin{array}{l} F(q_L)\left[u(w_{L,1})+\beta p_L u(w_{L,2})\right] \\ +(1-F(q_L))\left[u(z)+\beta p_L U_{1p}\right] \end{array} \right\} \\
\text{s.to} \quad & H(q)\left[y-w_{L,1}+\beta p_L(y-w_{L,2})\right]=C
\end{aligned}
$$

and

$$
\begin{aligned}
U_{1p} &= \max_{\{q,w\}} \left\{F(q)u(w)+(1-F(q))u(z)\right\} \\
\text{s.to} \quad & H(q)\left[y-w\right]=C
\end{aligned}
$$

We remark that the incentive constraint which prevents type H workers from misrepresenting their type as L is not binding hence not included, because we could verify that type H workers are worse off choosing the allocation of type L instead of sticking to their own.[14] This economy is an extension to the theoretical context proposed by Guerrieri, Shimer and Wright (2010). And the following result could be achieved:

Proposition 3. *In our economy, the separating equilibrium exists and is unique.*

Proof. To show the uniqueness and existence, it suffices to check that 3 mild assumptions proposed in Guerrieri, Shimer and Wright (2010), which are common in contract literature, are satisfied. In Appendix 1.6.3, the verification of these assumptions in our context is provided. □

Hence, we extend GSW's proof of the existence of a separating equilibrium to a two-period model. The interest is then to discuss the properties of this dynamic equilibrium with respect to observed features of the labor market where short and long term contracts coexist. We remark that the separating equilibrium always exists. The separating equilibrium here is immune to the Rothschild and Stigliz (1976) non-existence of equilibrium problem, because

[14] We refer the readers to proof of Proposition 4 for this.

with endogenous matching any pooling deviation could be proved to yield negative profit when the number of agents in the economy becomes large. We now turn to the properties of this equilibrium.

Proposition 4. *Under asymmetric information, the equilibrium allocation is such that:*

- *for type p_L, we have $q_L^{AI} = q_L^*$, and $w_{L,1}^{AI} = w_{L,2}^{AI} = w_L^*$;*
- *for type p_H, we have $w_{H,1}^{AI} < w_L^* < w_H^* < w_{H,2}^{AI}$;*
- *the type p_H workers are worse off in terms of expected utility, compared to the full-information case ($U_H^{AI} < U_H^*$);*
- *the type p_L worker receives the same expected utility ($U_L^{AI} = U_L^* < U_H^{AI}$);*
- *we have $q_H^* < q_H^{AI} < q_L^*$. Firms over-compensate workers in terms of total wage transfer compared to the full information case: $w_H^* - w_{H,1}^{AI} < \beta p_H \left(w_{H,2}^{AI} - w_H^*\right)$.*

Proof. See Appendix 1.6.4. □

In the presence of asymmetric information, separating equilibrium is achieved by increasing the wage of the type-H worker with his tenure in the firm. This wage profile discourages type-L workers from applying. Since type-L workers have a lower probability of staying, choosing this non-smoothed wage profile makes them receive a lower expected utility compared to the smoothed wage payment. In contrast, type-H workers have a higher probability of remaining at work tomorrow. Thus they are willing to sacrifice some wages today in exchange for higher earnings tomorrow.

Salop and Salop (1976) describe the same mechanism to induce self-selection. They consider a so-called Two-Part Wage (TPW) requiring that a new employee at the matching stage pays the firm an entrance fee of D_1, in return for which she will receive $w + D_2$ in the subsequent periods. In the equilibrium, the worker should pay her own turnover costs. This is equivalent to say that this worker guarantees that she will pay for the consequences of her potential premature quit. Our wage profile (Proposition 4) could also be interpreted in a fashion similar to Salop and Salop (1976). Recall that the free-entry condition in the presence of asymmetric information specifies $\frac{C}{H\left(q_H^{AI}\right)} = y - w_{H,1}^{AI} + \beta p_H \left(y - w_{H,2}^{AI}\right)$, and in the absence of asymmetric information $\frac{C}{H\left(q_H^*\right)} = y - w_H^* + \beta p_H \left(y - w_H^*\right)$. We could thus define the following difference: $D_1 = w_H^* - w_{H,1}^{AI} > 0$ and $D_2 = w_{H,2}^{AI} - w_H^* > 0$. Salop and Salop (1976) obtain the equality $\frac{D_1}{D_2} = \beta p_H$ (the self-selection ratio) at the equilibrium, due to worker's risk-neutrality. In our context, under risk-aversion, the result becomes $\frac{D_1}{D_2} < \beta p_H$ which is equivalent to $w_H^* + \beta p_H w_H^* < w_{H,1}^{AI} + \beta p_H w_{H,2}^{AI}$. Hence, firms have to over compensate workers in terms of wage transfer. On one hand, the 1st-period wage has to be reduced to a sufficiently low level to prevent undesirable types from applying. This reduction of 1st-period wage at the same time harms the good type's incentive from applying. On the other hand, the 2nd-period wage has to be set to a sufficiently high level to rescue the reduced incentives due to the decrease of 1st-period wage. Under risk aversion, the marginal gain from changing the 2nd-period wage is larger compared to changing the 1st-period wage, so that in order to attract workers through increasing 2nd-period wage, more than proportional amount of wage transfer should be foreseen.

To look into more details, let us turn to the first order conditions with respect to the wages. Notice that in case of asymmetric information, the first order conditions with respect to w_1 and w_2 are respectively

$$\begin{aligned} u'\left(w_{H,1}^{AI}\right) &= \frac{\lambda q_H^{AI}}{1-\delta} \\ u'\left(w_{H,2}^{AI}\right) &= \frac{\lambda q_H^{AI}}{1-\delta\frac{p_L}{p_H}} \end{aligned}$$

where λ is the Lagrangian multiplier for firms' breakeven constraint, and δ is the lagrangian multiplier associated with the incentive compatibility constraint. Dividing these two equations and rearranging, we are able to obtain

$$\frac{u'\left(w_{H,1}^{AI}\right)}{u'\left(w_{H,2}^{AI}\right)} = 1 + \frac{\delta}{1-\delta}\frac{p_H - p_L}{p_H}$$

The expression $\frac{\delta}{1-\delta}\frac{p_H-p_L}{p_H}$ could be regarded as a measure of the extent of agency cost in the economy. This index $\frac{\delta}{1-\delta}\frac{p_H-p_L}{p_H}$ is a product of δ, which measures the tightness of the IC constraint, and the likelihood ratio $\frac{p_H-p_L}{p_H}$, which measures the degree to which the agents' misbehavior (shirking) related to their realized type distorts the full information allocation. From the intuitive perspective, when $p_L = 0$, the information problem is the most serious in the economy, because the bad types will leave for sure in the subsequent period. This annihilates the firm's production. Thus, in this case, the incentive constraint should be "very" tight, with a "high" shadow price (cost). Hence, there should be a very large gap between w_1 and w_2 to induce some incentives. When $p_L \to p_H$, the IC problems becomes gradually less important (more relaxed), because the bad types behave almost as good as the good types. In this case, the IC constraint should be slack, with a low shadow price (cost). Hence, there should be gradually smaller gap between w_1 and w_2 to induce some incentives.

From the technical perspective, if $p_L \to 0$, then $\frac{\delta}{1-\delta}\frac{p_H - p_L}{p_H}$ is very large. Indeed, the likelihood ratio $\frac{p_H - p_L}{p_H}$ attains its maximal value 1. Moreover, the ratio $\frac{\delta(p_L)}{1-\delta(p_L)}$ also attains its largest value. Naturally, when the IC constraint is very tight, the shadow price δ (cost) is at its highest value. Hence, we have the largest gap between w_2 and w_1 when $p_L = 0$. If $p_L \to p_H$, then $\frac{\delta}{1-\delta}\frac{p_H - p_L}{p_H}$ will be very small. Firstly because $\frac{p_H - p_L}{p_H}$ tends to 0 in this case, and secondly because $\frac{\delta(p_L)}{1-\delta(p_L)}$ also takes its lowest value, because, in this case the IC constraint tends to be very slack, leading to a low value of the shadow price δ. Hence, when $p_L \to p_H$, there is no gap between w_1 and w_2, because there is no more information problem.

As a summary, risk aversion plays an important role in restricting the inter-temporal transfers. The following proposition examines to which extent the separating equilibrium is away from efficiency. Precisely, we compare the separating allocation with the pooling allocation, and we find that when the fraction of bad types is small enough, the separating allocation may be less efficient than the pooling allocation.[15]

Proposition 5. *The separating equilibrium may be Pareto inefficient under certain particular parameter values.*

Proof. See Appendix 1.6.5. □

If the share of the bad workers is small enough, the pooling allocation dominates the separating equilibrium, even though the pooling equilibrium never exists. Indeed, the optimization program of the pooling allocation approaches the optimization program for the type-H workers in the full information case, implying that $w_\alpha^* \to w_H^*$, $q_\alpha^* \to q_H^*$ and thus the good workers will be better off. Given that the allocation under full information is such that $q_H^* < q_L^*$ and $w_H^* > w_L^*$, the type L-workers could be employed more often, receive higher wages with this constructed pooling allocation and thus will also be better off.

1.3 Competitive search with endogenous separations

In the previous section, we have considered separation as an event which happens exogenously. In this section, we allow the agents to choose the period at which they quit the labor market. We study how this extensive margin can be manipulated by the optimal contract. The introduction of an outside option in the second part of the life cycle allows us to analyze how an incentive scheme distorts separation choices.[16] For simplicity, we assume that each

[15]In Salop and Salop (1976),there is no efficiency gain from introducing this self-selection mechanism, because both workers and firms are risk-neutral.

[16]If the outside option of the worker is an alternative job in another area, sector... the information of this job prospect is not directly observed by the employer. In the same spirit, if the outside option of the worker

worker is characterized by a specific realization of this outside option, which leads to a worker-specific tenure on this labor market.

In this context, a simple way to endogenize separation choice is to introduce an outside option which is worker-specific. Intuitively, low value of outside option induces the agent to stay, while high value of outside option induces the agent to leave the contract prematurely. In this sense, we say individuals' decisions about quitting are 'endogenous'.

We will distinguish two timings: in the first, each worker knows her characteristic before she searches for a job, whereas in the second, the information on her characteristic is revealed after the matching. In the first case, the asymmetry in the information is ex ante, whereas in the second, it is ex post.

1.3.1 Ex ante heterogeneity, quitting choice and dynamic contracts

We start with the case where the workers are ex ante different and know their value of second-period home production h before searching for a job. The value of h is assumed to be drawn from a distribution $G(h)$. According to the duration of the employment relationship, the labor market is segmented into two: the two-period, and the one-period submarket. We will abuse the terminology to call the two-period submarket the type-H market, and the one-period submarket the type-L market. Each worker, before searching for a job, has to choose which submarket to enter according to her value of h. There are several situations, and we now describe them heuristically. We start with full information.

Starting from workers with sufficiently small values of h. In case these workers do not find a job today, they will remain in the market and search again tomorrow, because their values of h are low compared to the value from search again. Then their choice is to enter the two-period (type-H) market. Using the same equilibrium definition as the previous section, the problem of such a worker with value h is written as follows[17]:

$$\begin{aligned} U^{*}_{H1}(h,z) \quad &= \quad \max_{\{q,w_{H,1},w_{H,2}\}} \left\{ \begin{array}{l} F(q_{H1})\left[u(w_{H1,1})+\beta u(w_{H1,2})\right] \\ +(1-F(q_{H1}))\left[u(z)+\beta U_{1p}(h,z)\right] \end{array} \right\} \\ \text{s.t.} \quad & H(q_{H1})\left[y-w_{H1,1}+\beta\left(y-w_{H1,2}\right)\right]=C \end{aligned}$$

is a non-participation value, linked to health, social security wealth or financial assets, firms must extract information by a mechanism design.

[17] Notice that we assumed h and z to be additively separable in terms of utility.

where

$$\begin{aligned} U_{1p}(h,z) &= \max_{\{q,w\}} \{F(q)u(w)+(1-F(q))[u(h)+u(z)]\} \\ \text{s.t.} \quad & H(q)[y-w]=C \end{aligned}$$

An important remark is that the posted wage and induced queue length both depend on h, although h does not enter explicitly into firms' profit equation. Indeed, workers of different h will have correspondingly different demands on wages, and there will always be firms which post corresponding wages to cater each workers' wage demand, as in Acemoglu and Shimer (1999): a competitive search context with unemployment insurance as outside option for workers. It could be verified by implicit function theorem that both the equilibrium wage and queue length are increasing in h. Technically, this two-period submarket is composed of a continuum of subsubmarket indexed by h.

Now we turn attention to the program $U_{1p}(h,z)$. We know that workers with high value of h will demand high wages corresponding to their value of h; and as h increases, there is a point of h above which the firm can no longer manage to attract workers. In that situation, those workers may not choose to search again tomorrow in case they do not secure a job today: they may simply leave the market and enjoy the home production $u(h)$ if $u(h)$ is such that $\max\{U_{1p}(h,z),u(h)\}=u(h)$. The corresponding program then becomes

$$\begin{aligned} U^*_{H2}(h,z) &= \max_{\{q,w_{H,1},w_{H,2}\}} \left\{ \begin{array}{l} F(q_{H2})[u(w_{H2,1})+\beta u(w_{H2,2})] \\ +(1-F(q_{H2}))[u(z)+\beta u(h)] \end{array} \right\} \\ \text{s.t.} \quad & H(q_{H2})[y-w_{H2,1}+\beta(y-w_{H2,2})]=C \end{aligned}$$

Now we are able to succinctly represent the above two situations in the following program:

$$\begin{aligned} U^*_{H}(h_H,z) &= \max_{\{q,w_{H,1},w_{H,2}\}} \left\{ \begin{array}{l} F(q_H)[u(w_{H,1})+\beta u(w_{H,2})] \\ +(1-F(q_H))\left[u(z)+\beta \max\left\{u(h_H),U_{1p}(h_H,z)\right\}\right] \end{array} \right\} \\ \text{s.t.} \quad & H(q_H)[y-w_{H,1}+\beta(y-w_{H,2})]=C \end{aligned}$$

Notice that in the above two situations, workers always have intention to work for two periods as long as they find the job today. The firms can observe workers' value of h. They could choose to either enter into the one-period or the two-period market, and they know that workers in the two-period market will never break the employment relationship unilaterally.

As h further increases, up to a point such that $u(h)$ is weakly greater than $u(w_{H,2})$, the corresponding worker becomes gradually unwilling to accept the job offer because she finds herself better off staying at home instead of working at the second period. To make the workers effectively unwilling to stay in the two-period market, it has to be that $u(h)-u(w_{H,2})$

is positive and sufficiently large, because if they choose the one-period market at the point $u(h) = u(w_{H,2})$, there would be a discrete drop in their expected utility, owing to our finding in the previous section that a two-period worker enjoys higher per period expected utility from search compared to a one-period worker. The firms know that workers with large values of h will prefer to stay at home at the second period (i.e. choose to enter the one-period market), so if they want to attract these workers, they should enter the one-period market. This justifies the existence of the one-period market at first period, because there is demand from workers' side. We write the problem for these potential one-period (type-L) workers' as follows:

$$
\begin{aligned}
U_L^*(h_L, z) \quad &= \quad \max_{\{q, w_L\}} \left\{ \begin{array}{l} F(q_L)\, u(w_L) + (1 - F(q_L))\, u(z) \\ +\beta u(h) \end{array} \right\} \\
\text{s.t.} \quad & \quad H(q_L)[y - w_L] = C
\end{aligned}
$$

Notice that these one-period workers do not have incentive to search for a job again in the second period. Suppose they have incentive to do so, then we shall substitute $u(h)$ for $U_{1p}(h,z)$. This implies that there is at some point $u(h)$ may be smaller than $U_{1p}(h,z)$. However, we have learnt from the above arguments that for the workers who enter into the one-period market, their level of h is such that $u(h) > u(w_{H,2}) > U_{1p}(h,z)$. We thus obtain a contradiction.

As a summary, a worker's decision on whether entering into a two-period (type-H) or a one-period (type-L) market is captured by the comparison of the expected utility between $U_H^*(h,z)$ and $U_L^*(h,z)$. If the h is such that $U_H^*(h,z) > U_L^*(h,z)$, this worker in question will choose to enter the two-period market; If the h is such that $U_H^*(h,z) < U_L^*(h,z)$, this worker chooses to enter into the one-period market; And if $U_H^*(h,z) = U_L^*(h,z)$ workers are indifferent, we assume that she always chooses to leave the labor market. We denote now $\hat{h}^{FI}$ as the threshold captured by $U_H^*\left(\hat{h}^{FI}, z\right) = U_L^*\left(\hat{h}^{FI}, z\right)$ in case of full information, and we regarded it as the separation threshold.

Equilibrium with dynamic contract

Our main focus in this section is to examine how does asymmetric information have impact on this separation threshold. To this end, we should check whether certain type with a given value of h can be better off by choosing the other types' optimal allocations, so that in case of asymmetric information there is potential misrepresentation of type. Since the value of h does not explicitly enter into firms' profit equation, it is straightforward to verify by revealed preference as in Acemoglu and Shimer (1999) that among the pool of workers with $h < \hat{h}^{FI}$ (within the two-period market), no worker will have incentive to choose to apply to wages destined for another worker of different h. As for the workers with $h \geq \hat{h}^{FI}$ (with in the one-period market), their allocation does not depend on h, so there is no information problem. Our problem is then reduced to whether the workers with $h \geq \hat{h}^{FI}$ (from two-period market) have incentive to misrepresent their types as certain workers with $h < \hat{h}^{FI}$ (from one-period market), as in the previous section. Recall that we term those workers with $h \geq \hat{h}^{FI}$ as type-H workers, and those workers with $h < \hat{h}^{FI}$ as type-H workers; We have the following result concerning workers' mimicking behavior.

Proposition 6. *Type-L workers are always better off by misrepresenting their types as type-H workers.*

Proof. See Appendix 1.6.6. □

This Proposition suggests that workers from one-period market indeed have intention to deviate to apply to wages posted in the two-period market. More precisely, given that the two-period market consists again of a continuum of submarkets indexed by h, workers from the one-period market will have incentive to misrepresent themselves to join all these submarkets. This result is expected. Notice that in the one-period market, the term associated with β becomes absent in the firm's profit equation. Due to this indirect impact from h, information problem arises between these two markets. Again, firms in the two-period market must implement a mechanism to avoid the risk of hiring a worker with a one-period (short) term on the job where the search costs can only be recouped by hiring a worker with a two-period (long) term on the job. Analogous to the previous section, the optimization program for a given h in the two-period market is written as follows:

$$
\begin{aligned}
U_H^{AI}(h_H, z) \quad &= \quad \max_{\{q, w_{H,1}, w_{H,2}\}} \left\{ \begin{array}{l} F(q_H)\left[u(w_{H,1}) + \beta u(w_{H,2})\right] \\ +\left(1 - F(q_H)\right)\left[u(z) + \beta \max\left\{u(h_H), U_{1p}(h_H, z)\right\}\right] \end{array} \right\} \\
\text{s.t.} \quad & \quad H(q_H)\left[y - w_{H,1} + \beta\left(y - w_{H,2}\right)\right] = C \\
IC \text{ constraint} \quad & \quad F(q_H)\left[u(w_{H,1})\right] + \left(1 - F(q_H)\right)\left[u(z_1)\right] + \beta u(h_L) \leq U_L
\end{aligned}
$$

where

$$
\begin{aligned}
U_{1p}(h, z) \quad &= \quad \max_{\{q, w\}} \left\{F(q)\, u(w) + (1 - F(q))\left[u(h) + u(z)\right]\right\} \\
\text{s.t.} \quad & \quad H(q)\left[y - w\right] = C
\end{aligned}
$$

The properties of the allocation of asymmetric information are summarized in the following proposition.

Proposition 7. *The optimal contract is such that*

- *for the wages,* $w_{H,1}^{AI} < w_L^* < w_H^* < w_{H,2}^{AI}$, *and* $w_{H,1}^{AI} + \beta w_{H,2}^{AI} > (1+\beta)w_H^*$.
- *for the application strategies (or the queue lengths),* $q_H^* < q_H^{AI} < q_L^{AI} = q_L^*$.

Proof. See Appendix 1.6.7. □

The incentive contract introduces an age-increasing wage. More precisely, by setting $w_{H,1}^{AI} < w_H^* < w_{H,2}^{AI}$, the agents who would choose to quit the labor market early are discouraged from applying. The mechanisms which lead to firms' over-compensation, i.e. $w_{H,1}^{AI} + \beta w_{H,2}^{AI} > (1+\beta)w_H^*$, share the same interpretation as the previous section. We move now to the next proposition on the separation threshold:

Proposition 8. *The threshold of the pooling equilibrium is such that* $\hat{h}^{AI} < \hat{h}^{FI}$.

Proof. By definition, we should have $U_H^{AI}(\hat{h}^{AI}) = U_L^{AI}(\hat{h}^{AI})$ and $U_H^*(\hat{h}^{FI}) = U_L^*(\hat{h}^{FI})$. Now suppose that $\hat{h}^{AI} \geq \hat{h}^{FI}$. This implies that $U_H^{AI}(\hat{h}^{FI}) \geq U_L^{AI}(\hat{h}^{FI})$, i.e., the type $\hat{h}^{FI}$ in the asymmetric information economy will not choose to enter into the one-period submarket. However, this in turn implies that $U_H^{AI}(\hat{h}^{FI}) \geq U_L^*(\hat{h}^{FI}) = U_H^*(\hat{h}^{FI})$, which is impossible, because we know $U_H^{AI} < U_H^*$ for any level of h. □

At the equilibrium, the optimal contract reaches a separating equilibrium while retaining fewer worker on long-term contract. This is a direct consequence of the separation decision. As in the previous context, to achieve separating equilibrium, it is necessary to distort the contract of type-H workers. However, here an additional loss is induced: the worker who should have remained active now chooses to enter into the one-period submarket. Indeed the firm must discourage the undesirable workers from applying for the job posted for the workers who want to work for two periods (the ones that prefer "long-term" contracts). However, this is done through firms' ex ante decisions in restricting the set of the workers who apply for long term contracts in order to screen the applicants.

By nature, when the risk from search is endogenously affected by the posted contract, the good workers have to pay for the information rent by receiving distorted allocation, which further decreases their expected utility, and leads them to quit in the second period. Hence, we observe that if the extensive margin (the separation decision) is 'endogenized' in this sense, the distorted wage offer operates through the extensive margin of search to lower the total labor supply when there is asymmetric information.

Importantly, this contrasts with the results of Guerrieri, Shimer and Wright (2010) where the separating contract leads to a rise of the intensive margin in the asymmetric information case. Nevertheless, in our dynamic context, even if the contract proposes an tenure-increasing wage, this cannot lead agents to prefer the long-term relationship because the screening process is done ex ante. This is not the case in Guerrieri, Shimer and Wright (2010) because the intensive margin can react immediately to the wage.

1.3.2 Ex post heterogeneity and credible commitment

In this section, we study the case where the workers learn their outside option value after getting matched with an employer. The difference between this economy and the previous ones is that now the heterogeneity is privately observed by the workers after the employment relationship has been established (ex post). Then the problem is no longer adverse selection, but one related to commitment: the contracts posted ex ante, if accepted by workers, should induce them to make proper ex post quitting decisions which are conformable to what is agreed ex ante.[18] In other words, the contract should ensure that ex post the workers have no incentive breaking it unilaterally. Specifically, we are in a situation where firms have full commitment on wage posting, whereas workers have limited commitment on premature quitting.

Our economy is closely related to Guerrieri (2008), who treats a competitive search context where workers hold private information on their idiosyncratic disutility realized immediately after the match is formed. Apart from difference in workers' preference and length of periods, we also differ in the fact that in Guerrieri (2008) workers' premature quitting incentive is driven by high disutility at work, while in our context it is driven by higher outside option.[19] How does workers' quitting incentive generate distortion? Whenever workers can quit prematurely (without commitment and without paying any extra cost), firms are not able to assure production and recoup the sunk vacancy cost. Anticipating this, firms want to propose a countermeasure which provides themselves more insurance.

We propose some straightforward modifications of the techniques developed in Guerrieri (2008) to solve our problem, and obtain the following results: *(i)* the equilibrium wage contract consists of an ascending wage profile where the wage per period is lower than the full information level; *(ii)* these low wages allow firms to assure higher ex post surplus, but lead workers less willing to stay due to comparatively more attractive outside option, so that a lower participation margin results, i.e. $\hat{h}^{AI} < \hat{h}^{FI}$ (qualitatively, this result is the same as

[18]With ex ante heterogeneity under asymmetric information, the employers' problem is to screen the undesirable types and to induce self-selection among the pool of workers knowing their characteristic before the employment relationship is established (ex ante).

[19]With this framework, we do not limit the heterogeneity to a match-specific event, but we also account for permanent distributions in outside opportunities that are available in different sectors/geographic areas or in preferences for inactivity.

the one obtained in the previous section), *(iii)* there is inferior job creation compared to the full information economy.

We will put emphasis on the elements different from the previous contexts. Workers and firms are now ex ante identical. Once a match is formed, the worker receives an outside option shock which takes place at the end of the 1st period during this employment relationship. Depending on the realization of the outside option, the worker decides to work for another period or to leave the market early (without paying any compensation to the firm). We assume that the level of the outside option shock h is drawn randomly from a cumulative distribution function $G(\cdot)$, within the interval $H \equiv \left[\underline{h}, \overline{h}\right]$. This realization is a private information of workers. Thus the firm's problem is to propose a wage contract that *(i)* guarantees participation (Individual rationality), and *(ii)* makes sure the workers reveal truthfully their home production realization (Incentive compatibility).

We consider direct revelation mechanisms, where a contract is a function which maps each reported h to a retaining probability $e(h) \in [0,1]$, as well as the wages $\omega(h) = (w_1(h), w_2(h)) \in (\mathbb{R}_+, \mathbb{R}_+)$. We use $J\left(h, \tilde{h}\right)$ to denote the ex post extra value from staying in the time between for the worker who is of type h but reports $\tilde{h}$, and it is written as:

$$J\left(h,\tilde{h}\right) = e\left(\tilde{h}\right) u\left(w_2\left(\tilde{h}\right)\right) - e\left(\tilde{h}\right) u(h)$$

Then incentive compatibility (IC) implies $J(h,h) \geq J\left(h,\tilde{h}\right)$ for all h and $\tilde{h}$ belonging to the support of h. Individual rationality (IR) requires $J(h,h) \geq 0$ for all h belonging to its support. According to the following lemma, we could reduce the number of the constraints related to IC and IR.

Lemma 1. *The IC and IR constraints are equivalent to the following conditions:*

(1) $e(h)$ *being non-increasing.*

(2) $J(h,h) - J\left(\overline{h},\overline{h}\right) = \int_h^{\overline{h}} e(x) u'(x) dx$

(3) $J\left(\overline{h},\overline{h}\right) \geq 0$

Proof. See Appendix 1.6.8. □

The competitive search equilibrium is adapted as follows:

Definition 2. *A competitive search equilibrium consists of a set of incentive compatible and individual rational mechanisms* $M = \{e(h), \omega(h)\}$ *for all h, an induced application strategy from workers* $Q^*(M)$*, as well as a utility function* U^{AI} *which satisfies:*

(1) Employers' profit maximization and free entry:

$$H\left(Q^{*}\left(M\right)\right)\times\left(y-w_{1}\left(h\right)+\beta\int_{\underline{h}}^{\overline{h}}e\left(h\right)\left(y-w_{2}\left(h\right)\right)dG\left(h\right)\right)\quad\leq\quad C$$
$$\textit{subject to IC and IR}$$

(2) Workers' optimal application:

$$\begin{aligned}U^{AI}\quad\geq\quad & F\left(Q^{*}\left(M\right)\right)\times\left[u\left(w_{1}\left(h\right)\right)+\beta\int_{\underline{h}}^{\overline{h}}\left\{\begin{array}{c}e\left(h\right)u\left(w_{2}\left(h\right)\right)\\+\left(1-e\left(h\right)\right)u\left(h\right)\end{array}\right\}dG\left(h\right)\right]\\ & +\left(1-F\left(Q^{*}\left(M\right)\right)\right)\times\left[u\left(z\right)+\beta\int_{\underline{h}}^{\overline{h}}\max\left\{U_{1p},u\left(h\right)\right\}dG\left(h\right)\right]\end{aligned}$$

where U^{AI} is given by

$$U^{AI}\quad=\quad\max_{M'}\left\{\begin{array}{c}F\left(Q^{*}\left(M'\right)\right)\times\left[u\left(w_{1}\left(h\right)\right)+\beta\int_{\underline{h}}^{\overline{h}}\left\{\begin{array}{c}e\left(h\right)u\left(w_{2}\left(h\right)\right)\\+\left(1-e\left(h\right)\right)u\left(h\right)\end{array}\right\}dG\left(h\right)\right]\\+\left(1-F\left(Q^{*}\left(M'\right)\right)\right)\times\left[u\left(z\right)+\beta\int_{\underline{h}}^{\overline{h}}\max\left\{U_{1p},u\left(h\right)\right\}dG\left(h\right)\right]\end{array}\right\}$$

In equilibrium, the functional relationship between the contract M and the queue length Q is learnt by both the workers and firms. From the firms' perspective, they anticipate how the posted incentive compatible and individually rational contract can induce the workers to choose their application strategy (captured by Q), even for contracts which are not posted in equilibrium. From the workers' perspective, given any wage contract, they optimally choose the contract among them, taking into account the trade-off between the contract and the probability of obtaining it - there is subgame perfection. At last, since the firms can anticipate that non-optimal contracts will not receive application (queue length is zero), they will never post such non-optimal contracts.

As in Guerrieri (2008), the competitive search equilibrium defined is equivalent to the solution of the following program:

$$\begin{aligned}U^{AI}\quad=\quad & \max_{\{e(h),w_{1}(h),w_{2}(h),q\}}\left\{\begin{array}{c}F\left(q\right)\times\left[u\left(w_{1}\left(h\right)\right)+\beta\int_{\underline{h}}^{\overline{h}}\left\{\begin{array}{c}e\left(h\right)u\left(w_{2}\left(h\right)\right)\\+\left(1-e\left(h\right)\right)u\left(h\right)\end{array}\right\}dG\left(h\right)\right]\\+\left(1-F\left(q\right)\right)\times\left[u\left(z\right)+\beta\int_{\underline{h}}^{\overline{h}}\max\left\{U_{1p},u\left(h\right)\right\}dG\left(h\right)\right]\end{array}\right\}\\ & \text{s.to}\quad e\left(h\right)\text{ being non-increasing}\\ & \qquad\text{IC, IR and Zero Profit condition.}\end{aligned}$$

With standard procedure in contract theory (see, e.g. Laffont and Martimort (2001)), and following Guerrieri (2008), we could further simplify the IC and IR conditions such that it suffices to solve the following program:

$$
\begin{aligned}
U^{AI} \quad = \quad & \max_{\{e(h),w_1(h),w_2(h),q\}} \left\{ \begin{array}{l} F(q) \times \left[u(w_1(h)) + \beta \int_{\underline{h}}^{\overline{h}} \left\{ \begin{array}{l} e(h)\, u(w_2(h)) \\ +(1-e(h))\, u(h) \end{array} \right\} dG(h) \right] \\ +(1-F(q)) \times \left[u(z) + \beta \int_{\underline{h}}^{\overline{h}} \max\left\{ U_{1p}, u(h) \right\} dG(h) \right] \end{array} \right\} \\
\text{s.to} \quad & e(h) \text{ being non-increasing} \\
& H(q) \times \left(y - w_1(h) + \beta \int_{\underline{h}}^{\overline{h}} e(h)(y - w_2(h))\, dG(h) \right) = C \\
& \int_{\underline{h}}^{\overline{h}} e(h) \left[u(w_2(h)) - u(h) - \frac{G(h)}{g(h)} u'(h) \right] dG(h) \geq 0
\end{aligned}
$$

The FOCs with respect to w_1 and w_2 show that wages are type-dependent, and increasing in time:

$$
\begin{aligned}
u'\left(w_1^{AI}\right) &= \lambda q^{AI} \\
u'\left(w_2^{AI}\right) &= \lambda q^{AI} \tfrac{1}{1+\delta}
\end{aligned}
$$

The pointwise maximization with respect to $e(h)$ yields

$$
u\left(w_2^{AI}\right) + u'\left(w_2^{AI}\right)\left(y - w_2^{AI}\right) \quad = \quad u\left(\hat{h}\right) + \tfrac{\delta}{1+\delta} \tfrac{G(\hat{h})}{g(\hat{h})} u'\left(\hat{h}\right)
$$

where the existence of $\hat{h}$ hinges on the term $\frac{G(\hat{h})}{g(\hat{h})} u'\left(\hat{h}\right)$. If without risk aversion, the monotone hazard rate property $d\frac{G(h)}{g(h)}/dh > 0$ is sufficient to guarantee a unique value of $\hat{h}$. With risk aversion, we notice that the increase is damped by the decrease of the $u'\left(\hat{h}\right)$. If $u'\left(\hat{h}\right)$ decreases faster than $\frac{G(\hat{h})}{g(\hat{h})}$ increases, the firm then is not able to induce self-selection with this mechanism, and there may be bunching. We abstract from the complexity through that direction and focus on the case where $\frac{G(\hat{h})}{g(\hat{h})} u'\left(\hat{h}\right)$ in total is increasing, so that a unique value of $\hat{h}$ is implied by this equality. In this case, it implies that this pointwise maximization leads to a cutoff value $\hat{h}$ such that

$$
e(h) = \begin{cases} 0 & \text{if } h \geq \hat{h} \\ 1 & \text{if } h < \hat{h} \end{cases}
$$

This is conformable to the restriction that $e(h)$ should be weakly decreasing, and we hence verify that the solution of the program is equivalent to the solution of the program where the monotonicity constraint on $e(h)$ is imposed.

The level of shadow value δ reflects to which extent the allocations are affected by asymmetric information. From the binding constraint associated with δ, we know that w_2^{AI} is increasing with $\hat{h}$. Together with the FOC with respect to $e(h)$, it implies that the higher the value of δ, the lower should be the value of threshold $\hat{h}$. In other words, when the constraint is tighter (measured by higher value of δ), in equilibrium less trade takes place (captured by lower value of $\hat{h}$).

In particular, when $\delta = 0$, the information constraint is not binding. Then, from the FOCs with respect to wages we deduce $\hat{w}_1^{FI} = \hat{w}_2^{FI}$; at the same time, it is straightforward to verify from the FOC with respect to $e(h)$ that $w_2^{FI} = y = \hat{h}^{FI}$. Putting these two equalities into the zero profit condition of firms, we find that the zero profit condition can only be satisfied when the fixed vacancy cost is nil ($C = 0$) as in Guerrieri (2008). Hence the economy is necessarily distorted away from the first best case as long as fixed vacancy cost is sunk ex ante and positive.

We now summarize in the following proposition the comparison of allocations between the full information and asymmetric information case:

Proposition 9. *In the asymmetric information case, ie. when $\delta > 0$, we have*

- $w_1^{AI} < w_2^{AI} < y = w^{FI}$,
- $\hat{h}^{AI} < \hat{h}^{FI}$ *and*
- $q^{AI} > q^{FI}$

Since $q^{AI} > q^{FI}$ implies $F(q^{FI}) > F(q^{AI})$, and since $G(\hat{h}^{FI}) > G(\hat{h}^{AI})$, we are able to conclude that $F(q^{FI})G(\hat{h}^{FI}) > F(q_{AI})G(\hat{h}_{AI})$, which implies that the job matching is inferior in the asymmetric information case.

Proof. See Appendix 1.6.9 □

This proposition shows that an increasing wage profile is induced in order to achieve commitment, contrary to the case with full information where the wages are flat. Given that the reservation threshold of the outside option is lower in the economy with asymmetric information, there is less trade in the economy with asymmetric information, in other words, there is under-employment.

Notice that we have established that in both the case of ex ante and the case of ex post heterogeneity, this extensive margin is downwards distorted, however, the driving forces are not the same. With ex ante heterogeneity, firms can post wage profiles to induce self-selection by preventing the undesired workers from applying. With ex post heterogeneity, no one has information priority before search. The problem of the firm is to design a mechanism which consists of the wage profile, together with a rationing rule, such that workers truthfully reveal their type, based on which firms decide whether to retain the worker in question. This costly ex post information rent implies a discrete drop of ex post surplus for the firms, which hence reduces firms' ex ante incentive of vacancy opening - less firms enter and the workers/firms ratio increases: workers become difficult to get matched, and there is less job matching from the workers' perspective.

1.4 Extension: A three-period framework

We have studied a two-period model in the benchmark, where workers when facing a short-term separation shock, no longer remain in the market. Such a separation shock is absorbing, in the sense that whenever workers are "shocked" out of the labor force once, the economy ends. In this section, we propose a straightforward extension of the benchmark which allows us to account for the transitional aspect of the labor market occupations/states. A sufficient way is to extend our two-period model to a three-period economy, where workers face recurrently separation shock, and those who are shocked in the second period have another chance of re-entering into the market. We will solve this straightforward extension, and discuss possible future developments upon it.

With a little abuse of terminology, we will use the subscript l (long-term contract) for the workers who have potentially three periods of working horizon, the subscript m (median-term contract) for those with two periods of horizon, and the subscript s (short-term contract) for those with one period of horizon. The parameter p still represents workers' probability to keep her job. Extending the definition of competitive search equilibrium to three periods, we arrive at the following maximization program with full information:

$$\begin{array}{lll} U_{l,1} & = & \max\limits_{\{w_{l,1},w_{l,2},w_{l,3},q_{l,1}\}} \left\{ \begin{array}{c} F\left(q_{l,1}\right)\left[u\left(w_{l,1}\right)+\beta\left(\begin{array}{c} p\times\left[u\left(w_{l,2}\right)+\beta pu\left(w_{l,3}\right)\right] \\ +(1-p)\times\beta pU_{l,3}\end{array}\right)\right] \\ +\left(1-F\left(q_{l,1}\right)\right)\left[u\left(z\right)+\beta pU_{l,2}+\beta\left(1-p\right)\times\beta pU_{l,3}\right] \end{array}\right\} \\ \text{s.to} & & H\left(q_{l,1}\right)\left[y-w_{l,1}+\beta p\left(y-w_{l,2}\right)+\left(\beta p\right)^{2}\left(y-w_{l,3}\right)\right]=C \end{array}$$

where $U_{l,2}=U_{m,1}$, given that the three-period workers at the second period has the same behavior as a two-period worker,

$$\begin{array}{lll} U_{m,1} & = & \max\limits_{\{w_{m,1},w_{m,2},q_{m,1}\}} \left\{ \begin{array}{c} F\left(q_{m,1}\right)\left[u\left(w_{m,1}\right)+\beta pu\left(w_{m,2}\right)\right] \\ +\left(1-F\left(q_{m,1}\right)\right)\left[u\left(z\right)+\beta pU_{m,2}\right] \end{array}\right\} \\ \text{s.to} & & H\left(q_{m,1}\right)\left[y-w_{m,1}+\beta p\left(y-w_{m,2}\right)\right]=C \end{array}$$

And based on similar argument, we have $U_{l,3} = U_{m,2} = U_{s,1}$, for the last period of three-period workers,

$$\begin{array}{rcl} U_{s,1} & = & \max\limits_{\{w_{s,1},q_{s,1}\}} \left\{ \begin{array}{c} F(q_{s,1})[u(w_{s,1})] \\ +(1-F(q_{s,1}))[u(z)] \end{array} \right\} \\ \text{s.to} & & H(q_{s,1})[y-w_{s,1}] = C \end{array}$$

In section 2, we have shown that the per-period wage $w_{m,1}^{FI} = w_{m,2}^{FI} = w_m^{FI}$ is increasing, while the expected queue length q_m^{FI} decreasing in β and p. So that we conclude $w_m^{FI} > w_s^{FI}$, and $q_m^{FI} < q_s^{FI}$. We can generalize this result using the same method in the proof of proposition 1 that *"The per-period wage $w_{l,1}^{FI} = w_{l,2}^{FI} = w_{l,3}^{FI} = w_l^{FI}$ is increasing, while the expected queue length q_l^{FI} is decreasing in β and p"*. By this, we could conclude $w_l^{FI} > w_s^{FI}$, and $q_l^{FI} < q_s^{FI}$. The result that $w_l^{FI} > w_m^{FI}$ and $q_l^{FI} < q_m^{FI}$ can also be obtained using the same argument: simply by introducing another parameter k homogenous to the scalar $(\beta p)^2$, such that when $k = 0$, it coincides with $U_{m,1}$, and when $k = 1$, it coincides with $U_{l,1}$.

Now we move to the case where information is asymmetric. We assume that firms are able to perfectly observe the workers' working horizon ($i \in \{s,m,l\}$), and can post wages conditional on it. Then there is no information problem with respect to it. However, there is unobservable heterogeneity $p \in \{p_L, p_H\}$ within each working horizon, i.e., $p_H > p_L > 0$. In the previous section, we have dealt with the case for two period workers (median-term). Here we examine the case for three-period workers (long-term). To avoid multi-dimensional asymmetric information, we focus on the case where workers are either type L (with p_L) or type H (with p_H) during her employment expected lifespan.[20]

The corresponding program under asymmetric information is as follows:

$$\begin{array}{rcl} U_{l,1}^{AI}(H) & = & \max\limits_{\{w_{l,1},w_{l,2},w_{l,3},q_{l,1}|H\}} \left\{ \begin{array}{c} F(q_{l,1}) \left[u(w_{l,1}) + \beta \left(\begin{array}{c} p_H \times [u(w_{l,2}) + \beta p_H u(w_{l,3})] \\ +(1-p_H) \times \beta p_H U_{l,3} \end{array} \right) \right] \\ +(1-F(q_{l,1})) \left[u(z) + \beta p_H U_{l,2} + \beta(1-p_H) \times \beta p_H U_{l,3} \right] \end{array} \right\} \\ \text{s.to} & & H(q_{l,1}) \left[y - w_{l,1} + \beta p_H (y - w_{l,2}) + (\beta p_H)^2 (y - w_{l,3}) \right] = C \\ & & \left\{ \begin{array}{c} F(q_{l,1}) \left[u(w_{l,1}) + \beta \left(\begin{array}{c} p_L \times [u(w_{l,2}) + \beta p_L u(w_{l,3})] \\ +(1-p_L) \times \beta p_L U_{l,3} \end{array} \right) \right] \\ +(1-F(q_{l,1})) \left[u(z) + \beta p_L U_{l,2} + \beta(1-p_L) \times \beta p_L U_{l,3} \right] \end{array} \right\} \leq U_{l,1}(L) \end{array}$$

where

[20] This assumption is not in contradiction with the empirical fact that the hazard rate of separations declines with the tenure (see e.g. Farber (1994)). Indeed, the share of the "bad-type" workers declines in the population of worker having a long tenure. Hence, even with a constant separation rate during the contract, the heterogeneity of the worker population can lead to a decreasing relation between separation and tenure.

$$
\begin{array}{lcl}
U_{l,1}(L) & = & \max\limits_{\{w_{l,1},w_{l,2},w_{l,3},q_{l,1}|L\}} \left\{ \begin{array}{c} F(q_{l,1}) \left[u(w_{l,1}) + \beta \left(\begin{array}{c} p_L \times [u(w_{l,2}) + \beta p_L u(w_{l,3})] \\ +(1-p_L) \times \beta p_L U_{l,3} \end{array} \right) \right] \\ +(1-F(q_{l,1}))\left[u(z) + \beta p_L U_{l,2} + \beta(1-p_L) \times \beta p_L U_{l,3}\right] \end{array} \right\} \\
\text{s.to} & & H(q_{l,1}) \left[y - w_{l,1} + \beta p_L (y - w_{l,2}) + (\beta p_L)^2 (y - w_{l,3}) \right] = C
\end{array}
$$

Wage growth and tenure. With exactly the same argument, we can replicate the results that for the expected number of applicants we have $q_l^{FI}(H) < q_l^{AI}(H) < q_l^{AI}(H) = q_l^{FI}(L)$, and for the wages we have $w_{l,1}^{AI}(H) < w_{l,2}^{AI}(H) < w_{l,3}^{AI}(H)$, and $w_{l,1}^{AI}(H) < w_l^{FI}(H) < w_{l,3}^{AI}(H)$. Most importantly, with three periods, we are able to provide the profile of the wage during the worker' career in a firm, i.e. the link between the wage growth and the tenure, conditional on the worker's type.

Proposition 10. *When $u^{'''}(.) < 0$, the wage growth of the type H worker is concave: $w_{l,2}^{AI}(H) - w_{l,1}^{AI}(H) < w_{l,3}^{AI}(H) - w_{l,2}^{AI}(H)$*

Proof. See Appendix 1.6.10 □

Firstly, our results give some foundation to the observed increase of the real wage with tenure.[21] Secondly, our model generates a composition effect in the measure of tenure on the wage growth. Indeed, heterogeneous workers ("bad type" and "good type") can have the same tenure but not the same contract: "good type" workers have an increasing wage (Proposition 10), whereas "bad type" workers have a flat wage. This explains why the "staying" workers have larger wage increases linked to tenure than "moving" workers (see Altonji and Williams (1998)).

Tenure-dependent probability of separation. In the above extension of the baseline economy, we maintained the homogeneity of the separation probability p along the contract duration to avoid difficulties arising from multi-dimensional heterogeneity. However, there are good reasons that the separation probability p may be varying along the duration of the contract.

Firstly, p may be tenure-decreasing: workers with higher tenure can be closer to retirement, and hence have higher probability of leaving the labor market. Our model could partially capture this scenario. Consider the following example. Instead of having tenure-independent p, we now distinguish the second-period and third-period p as "$p_2(H)$ and

[21]Even if the tenure' contribution is not the largest component of the wage growth, all empirical studies show that it is significant and concave: see Altonji and Williams (1998) or Bagger, Fontaine, Postel-Vinay and Robin (2014).

$p_3(H)$" for type H, while "$p_2(L)$ and $p_3(L)$" for type L.[22] Assume further $\frac{p_2(L)}{p_2(H)} \leq \frac{p_3(L)}{p_3(H)} \leq 1$, where at least one inequality should hold strict to guarantee heterogeneity. These inequalities mean that at the end of the contract (period 3), the probability for workers of different types differs to a less extent. Then our above results on the queue length and wage profile can carry through, and the proof is exactly the same as those in section 2.[23]

Another situation arises when p depends positively on workers' tenure. In this case, one can preserve our results on the queue length and on the wage profile only if we introduce some restrictions on the parameters ensuring that the increase in p is not "too large".

The unemployment benefits. At last, we would like to raise some concern on the different effects from the working horizon (measured by p), and the outside option (measured by z), which can include unemployment benefits. It is well known from Acemoglu and Shimer (1999) that a higher outside option z leads risk-averse workers to demand higher wages more aggressively. This implies that implicitly we have both $\frac{dw}{dz} > 0$ and $\frac{dq}{dz} > 0$. While recall that we have shown $\frac{dw}{dp} > 0$ but $\frac{dq}{dp} < 0$, albeit higher value of p actually implies higher outside option. It is thus also important to study separately these two effects, and give conditions under which the working horizon (or outside option) effect dominates. This will be left for future work, and the comparative statics offered in the proof of Proposition 1 will serve as a useful tool.

1.5 Conclusion

In this paper, we deal with an asymmetric information problem within competitive search environment. Firms sink investment in vacancies ex ante and post wages in the hope of attracting the right workers to recoup invest costs; however, whether to quit the employment relationship prematurely is worker's private information; hence the posted wage profiles should induce self-selection.

In the baseline model where separation is an exogenous event, we firstly identify that it is the 'bad' workers (those with shorter horizon) who have incentive to misrepresent their type by applying for a long-term contract. We show that the optimal separating contract results in a distorted allocation for the 'good' workers (those with longer horizon) in order to prevent the 'bad' from applying for these jobs. The equilibrium wage profile has to be increasing, albeit premature quitting has the same consequence as a productivity drop. Our solution

[22] We still assume there are only two types to avoid multi-dimensional heterogeneity.

[23] Remark that another scenario implying a tenure-decreasing p is a case where past experiences of frequently unemployment episodes may discourage workers to search for an alternative opportunities (loss of trust, stigmatization effect...).

is close to the one discussed by Salop and Salop (1976), who considers a Two-Part Wage mechanism to induce self-selection.[24] We generalize the Salop and Salop (1976) result in a context of an equilibrium search model where risk aversion provides sufficient restrictions for the uniqueness of wage path and we extend the Guerrieri, Shimer and Wright (2010) result with wage dynamics and verify that this contract ensures the existence of a separating equilibrium. Finally, it is straightforward to observe that when the fraction of bad types is low, a pooling allocation can be constructed to make both types of workers better off, so that the separating equilibrium could be Pareto dominated.

In the second part, the separation decision is endogenized by introducing home production value as the source of asymmetric information. Such an extension permits us to study the impact of asymmetric information on the extensive margin. Two cases are considered: ex ante or ex post heterogeneity. For the case with ex ante heterogeneity, we obtain the same increasing wage profile which is necessary to achieve self-selection, as in the benchmark case where separation is exogenous. In addition, the extensive margin is distorted downwards, in the sense that more workers switch to search in the one-period submarket, and choose to be inactive in the second period. This is because the contract posted by the firms in the two-period is distorted for the sake of discouraging the one-period workers from applying, and this in turn leads to lower anticipated expected payoff in the two-period submarket. Finally, in the case of ex-post heterogeneity (e.g. in the case of an unexpected outside option shock during the career), since asymmetric information takes place after matching, there is ex ante only one submarket. The mechanism comprises a wage profile, as well as a retaining decision; in equilibrium, the wage profile is increasing, while the retaining decision leads to a cut-off value on the realisation of home production level, above which the workers in question choose to quit prematurely. The extensive margin compared to the full information economy is lower due to the existence of costly information rent. Thus, it seems that excessive separation linked to the asymmetry in the information is a robust result in a competitive search environment.

[24] This Two-Part Wage device requires that a new employee pays the firm an entrance fee, in return for which she will receive in the subsequent periods a higher wage (the wages increase with the tenure of the workers). In the equilibrium, the workers should pay for their own turnover costs. This is equivalent to saying that this worker pays a deposit which is a guarantee corresponding to the consequences of the lack of commitment.

1.6 Appendix

1.6.1 Proof of the Proposition 1

The solution for a type p_H worker is determined by the following two equations [25] (By substituting L for H, we obtain the solution for the type p_L workers):

$$\begin{aligned} \frac{1-\varepsilon_H(q_H^*)}{\varepsilon_H(q_H^*)} &= \frac{y-w_H^*}{\frac{1}{u'(w_H^*)}\times\left[u(w_H^*)-\frac{1}{1+\beta p_H}\left(\beta p_H U_{1p}+u(z)\right)\right]} \\ C &= H(q_H)(1+\beta p_H)(y-w_H) \end{aligned}$$

Result 1: $w_H^* > w_L^*$. Since type H's problem coincides with the type L's problem when $p_H = p_L$, if we could verify that $\frac{dw_H^*}{dp_H} > 0$, it implies that $w_H^* > w_L^*$. We consider the implicit function defined by rearranging the first order condition:

$$I(w_H^*, p_H) = u(w_H^*) - \frac{1}{1+\beta p_H}\left(\beta p_H U_{1p} + u(z)\right) - u'(w_H^*)(y-w_H^*)\frac{\varepsilon_H(q_H^*(w_H^*))}{1-\varepsilon_H(q_H^*(w_H^*))} = 0$$

where $q_H^*(w_H^*)$ is defined by the Free-entry condition. We want to show $\frac{dw_H^*}{dp_H} = -\frac{I_{p_H}}{I_{w_H^*}} > 0$, where $I_{p_H} = \frac{dI}{dp_H}$ and $I_{w_H^*} = \frac{dI}{dw_H^*}$. Given $I(w_H^*, p_H)$, we have

$$\begin{aligned} I_{p_H} = {} & \underbrace{-\frac{\beta}{(1+\beta p_H)^2}U_{1p} + \frac{\beta}{(1+\beta p_H)^2}u(z)}_{<0} \\ & \underbrace{-u'(w_H^*)(y-w_H^*)\frac{1}{(1-\varepsilon_H)^2}\varepsilon_H'\frac{-\beta C}{(1+\beta p_H)^2(y-w_H^*)H'(q_H^*)}}_{<0} < 0 \end{aligned}$$

On the other hand, $I(w_H^*, p_H)$ leads to

$$\begin{aligned} I_{w_H^*} = {} & \underbrace{u'(w_H^*)}_{>0}\underbrace{-u''(w_H^*)(y-w_H^*)\frac{\varepsilon_H(q_H^*(w_H^*))}{1-\varepsilon_H(q_H^*(w_H^*))}}_{>0}\underbrace{+u'(w_H^*)\frac{\varepsilon_H(q_H^*(w_H^*))}{1-\varepsilon_H(q_H^*(w_H^*))}}_{>0} \\ & \underbrace{-u'(w_H^*)(y-w_H^*)\frac{\varepsilon_H'}{(1-\varepsilon_H)^2}\times\frac{1}{H'(q_H^*)}\frac{C}{(1+\beta p_H)}\frac{1}{(y-w_H^*)^2}}_{>0} \end{aligned}$$

Thus in total $\frac{dw_H^*}{dp_H} > 0$, leading to $w_H^* > w_L^*$.

[25] The elasticity for the workers' job finding rate and the elasticity for the firms' hiring probability admit the following relationship: $\varepsilon_F(q_p) = \varepsilon_H(q_p) - 1$.

Result 2: $q_H^* < q_L^*$. The simplest way is to use the same argument as above and to prove $\frac{dq_H^*}{dp_H} < 0$. We thus consider the implicit function defined and obtained by rearranging the first order condition:

$$\begin{aligned}
X\left(q_H^*, p_H\right) &= 0 \\
&= u\left(w_H^*\left(q_H^*\right)\right) - \tfrac{1}{1+\beta p_H}\left(\beta p_H U_{1p} + u(z)\right) - u'\left(w_H^*\left(q_H^*\right)\right)\left(y - w_H^*\left(q_H^*\right)\right)\tfrac{\varepsilon_H\left(q_H^*\right)}{1-\varepsilon_H\left(q_H^*\right)}
\end{aligned}$$

where $w_H^*\left(q_H^*\right)$ is defined by the Free-entry condition. Implicit function theorem leads to $\frac{dq_H^*}{dp_H} = -\frac{X_{p_H}}{X_{q_H^*}}$, where $X_{q_H^*} = \frac{dX}{dq_H^*}$ and $X_{p_H} = \frac{dX}{dp_H}$. These derivatives are given by:

$$\begin{aligned}
X_{q_H^*} &= \underbrace{u' \times \frac{dw_H^*}{dq_H^*}}_{>0} \underbrace{- u'' \times \frac{dw_H^*}{dq_H^*}\left(y - w_H^*\left(q_H^*\right)\right)\frac{\varepsilon_H\left(q_H^*\right)}{1-\varepsilon_H\left(q_H^*\right)}}_{<0} \\
&\quad \underbrace{+ u' \times \frac{dw_H^*}{dq_H^*} \times \frac{\varepsilon_H\left(q_H^*\right)}{1-\varepsilon_H\left(q_H^*\right)}}_{>0} \underbrace{- u' \times \left(y - w_H^*\left(q_H^*\right)\right)\frac{1}{\left(1-\varepsilon_H\right)^2}\varepsilon_H'}_{<0} \\
&> 0
\end{aligned}$$

$$\begin{aligned}
X_{p_H} &= \overbrace{\underbrace{u' \times \frac{dw_H^*}{dp_H}}_{>0} \underbrace{- \frac{\beta}{\left(1+\beta p_H\right)^2}U_{1p} + \frac{\beta}{\left(1+\beta p_H\right)^2}u(z)}_{<0}}^{\gtrless 0} \\
&\quad \underbrace{- u'' \times \frac{dw_H^*}{dp_H}\left(y - w_H^*\left(q_H^*\right)\right)\frac{\varepsilon_H\left(q_H^*\right)}{1-\varepsilon_H\left(q_H^*\right)}}_{<0} \underbrace{+ u' \times \frac{dw_H^*}{dp_H} \times \frac{\varepsilon_H\left(q_H^*\right)}{1-\varepsilon_H\left(q_H^*\right)}}_{>0}
\end{aligned}$$

There seems to be ambiguity for the sign of X_{p_H}. We now try to rearrange it:

$$\begin{aligned}
X_{p_H} &= u' \times \tfrac{dw_H^*}{dp_H} - \tfrac{\beta}{\left(1+\beta p_H\right)^2}U_{1p} + \tfrac{\beta}{\left(1+\beta p_H\right)^2}u(z) \\
&\quad - u'' \times \tfrac{dw_H^*}{dp_H}\left(y - w_H^*\left(q_H^*\right)\right)\tfrac{\varepsilon_H\left(q_H^*\right)}{1-\varepsilon_H\left(q_H^*\right)} + u' \times \tfrac{dw_H^*}{dp_H} \times \tfrac{\varepsilon_H\left(q_H^*\right)}{1-\varepsilon_H\left(q_H^*\right)} \\
&= \tfrac{\beta}{\left(1+\beta p_H\right)^2} \times \left[u' \tfrac{1}{1-\varepsilon_H}\tfrac{C}{H\left(q_H^*\right)} - \left(U_{1p} - u(z)\right)\right] \\
&\quad + \tfrac{\beta}{\left(1+\beta p_H\right)^2} \times \left[-u'' \times \left(y - w_H^*\right)\tfrac{\varepsilon_H}{1-\varepsilon_H}\tfrac{C}{H\left(q_H^*\right)}\right]
\end{aligned}$$

We use first order condition to rewrite it, and obtain

$$\begin{aligned}
X_{p_H} &= \frac{\beta}{(1+\beta p_H)^2} \times \left[u' \frac{1}{1-\varepsilon_H} \frac{C}{H(q_H^*)} - (U_{1p} - u(z)) - u'' \times (y - w_H^*) \frac{\varepsilon_H}{1-\varepsilon_H} \frac{C}{H(q_H^*)}\right] \\
&= \frac{\beta}{(1+\beta p_H)^2} \times \left[\overbrace{\frac{1}{\varepsilon_H} [\beta p_H u(w_H^*) - \beta p_H U_{1p}]}^{>0} + \underbrace{\overbrace{\frac{1}{\varepsilon_H}}^{>1} [u(w_H^*) - u(z)] - (U_{1p} - u(z))}_{>0}\right] \\
&\quad + \frac{\beta}{(1+\beta p_H)^2} \times \left[-u'' \times (y - w_H^*) \frac{\varepsilon_H}{1-\varepsilon_H} \frac{C}{H(q_H^*)}\right] \\
&> 0
\end{aligned}$$

We thus deduce that $X_{p_H} > 0$, and thus $\frac{dq_H^*}{dp_H} = -\frac{X_{p_H}}{X_{q_H^*}} < 0$. This implies that the surplus of the firms is increasing in the value of p_H, leading to $q_H^* < q_L^*$.

Result 3: $U_H^* > U_L^*$. When $p_L = p_H$, $U_L^* = U_H^*$. In addition, since by Envelope Theorem the U_L^* is increasing in p_L (higher value of p corresponds to higher value from search), we have $U_H^* > U_L^*$.

Result 4: $(1+\beta p_H) w_H^* - (1+\beta p_L) w_L^* < (1+\beta p_H) y - (1+\beta p_L) y$. By the binding zero profit condition for both types, we have

$$\begin{aligned}
(1+\beta p_H) y - (1+\beta p_H) w_H^* &= \frac{C}{H(q_H^*)} \\
(1+\beta p_L) y - (1+\beta p_H) w_L^* &= \frac{C}{H(q_L^*)}
\end{aligned}$$

Since $q_H^* < q_L^*$, we have $(1+\beta p_H) w_H^* - (1+\beta p_L) w_L^* < (1+\beta p_H) y - (1+\beta p_L) y$.

1.6.2 Proof of Proposition 2

Proof. When $p_H = 1$ and $p_L = 0$, the programs are reduced to

$$\begin{aligned}
U_H^* &= F(q_H^*) \left[u\left(w_{H,1}^*\right) + \beta u\left(w_{H,2}^*\right)\right] + (1 - F(q_H^*)) \left[u(z) + \beta U_{1p}\right] \\
U_L^* &= F(q_L^*) [u(w_L^*)] + (1 - F(q_L^*)) [u(z)]
\end{aligned}$$

We need to show that $U_H^* > (1+\beta) U_L^*$. Recall that when $p_L = 0$, $U_L^* = U_{1p}$. We could reorganize terms in the following manner:

$$\begin{aligned}
U_H^* &= F(q_H^*) \left[u\left(w_{H,1}^*\right) + \beta u\left(w_{H,2}^*\right)\right] + (1 - F(q_H^*)) \left[u(z) + \beta U_{1p}\right] \\
&= F(q_H^*) \left(u\left(w_{H,1}^*\right) - u(z)\right) + u(z) + \beta \left[F(q_H^*) \left(u\left(w_{H,1}^*\right) - U_{1p}\right) + U_{1p}\right] \\
&> U_L^* + \beta \left[F(q_H^*) \left(u\left(w_{H,1}^*\right) - U_{1p}\right) + U_{1p}\right] \\
&> U_L^* + \beta U_L^*
\end{aligned}$$

where the first inequality is induced by virtue of Proposition 1.

1.6.3 Proof of Proposition 3

A separating contract always exists. We will show that the equilibrium allocation satisfies the 3 main assumptions (emphasized by GSW), which are required for the existence and uniqueness of the competitive search equilibrium.

A1) Monotonicity. Given any wage profile (w_1, w_2), the ex post surplus of firms should be monotonic in p. This is satisfied, because $(y-w_1)+\beta p_H(y-w_2) > (y-w_1)+\beta p_L(y-w_2)$

A2) Local non-satiation. Let $(w_{i,1}, w_{i,2})$ be a feasible contract for type i. Consider $\hat{w}_1 = w_{i,1}+h$, and $\hat{w}_2 = w_{i,2}+k$ in the neighborhood of $(w_{i,1}, w_{i,2})$, so that $y-w_{i,1}+\beta p_i(y-w_{i,2})-(h+\beta p_i k)\geq 0$, and $\sqrt{h^2+k^2}<\varepsilon$. Then we require

(1) $y-\hat{w}_1+\beta p_i(y-\hat{w}_2) > y-w_{i,1}+\beta p_i(y-w_{i,2})\geq 0$, and

(2) for $j<i$ $(p_j<p_i)$,

$u(\hat{w}_1)+\beta p_j u(\hat{w}_2)-u(z)-\beta p_j U_{1p}\leq u(w_{i,1})+\beta p_j u(w_{i,2})-u(z)-\beta p_j U_{1p}$.

It suffices that we have (1) $h+\beta p_i k<0$, and (2) $u^{'}(w_{i,1})h+\beta p_j u^{'}(w_{i,2})k\leq 0$. Or equivalently, $\min\left\{-\beta p_i, -\beta p_j\frac{u^{'}(w_{i,2})}{u^{'}(w_{i,1})}\right\} > \frac{h}{k}$ for $h<0$ and $k>0$.

A3) Sorting (similar to single-crossing). Given any type i, and given any feasible wage profile $(w_{i,1}, w_{i,2})$ for this type i, there exists a wage profile $(\hat{w}_1, \hat{w}_2)$ which is close enough to $(w_{i,1}, w_{i,2})$, i.e., in some ε neighborhood of $(w_{i,1}, w_{i,2})$ such that

1) for $j\geq i$ (types better than i):

$$u(\hat{w}_1)+\beta p_j u(\hat{w}_2)-u(z)-\beta p_j U_{1p} \quad \geq \quad u(w_{i,1})+\beta p_j u(w_{i,2})-u(z)-\beta p_j U_{1p}$$

2) for $j<i$ (types worse than i):

$$u(\hat{w}_1)+\beta p_j u(\hat{w}_2)-u(z)-\beta p_j U_{1p} \quad < \quad u(w_{i,1})+\beta p u(w_{i,2})-u(z)-\beta p_j U_{1p}$$

Proof of A3): Let $(w_{i,1}, w_{i,2})$ be a feasible contract for type i. Consider $\hat{w}_1 = w_{i,1}+h$, and $\hat{w}_2 = w_{i,2}+k$ in the neighborhood of $(w_{i,1}, w_{i,2})$, so that $y-w_{i,1}+\beta p_i(y-w_{i,2})-(h+\beta p_i k)\geq 0$, and $\sqrt{h^2+k^2}<\varepsilon$.

For $j\geq i$, we require

$u(\hat{w}_1)+\beta p_j u(\hat{w}_2)-u(z)-\beta p_j U_{1p} > u(w_{i,1})+\beta p_j u(w_{i,2})-u(z)-\beta p_j U_{1p}$

For $j^{'}<i$, we require

$$u(\hat{w}_1) + \beta p_{j'} u(\hat{w}_2) - u(z) - \beta p_{j'} U_{1p} < u(w_{i,1}) + \beta p_{j'} u(w_{i,2}) - u(z) - \beta p_{j'} U_{1p}$$

The former can be satisfied as long as $u'(w_{i,1})h + \beta p_j u'(w_{i,2})k > 0$; The latter can be satisfied as long as $u'(w_{i,1})h + \beta p_{j'} u'(w_{i,2})k < 0$. Equivalently, we should require $-\beta p_{j'} \frac{u'(w_{i,2})}{u'(w_{i,1})} > \frac{h}{k} > -\beta p_j \frac{u'(w_{i,2})}{u'(w_{i,1})}$ with $h < 0$ and $k > 0$, which is possible for $p_{j'} < p_j$.

1.6.4 Proof of Proposition 4

Claim 1: $w_{H,1}^{AI} < w_{H,2}^{AI}$. Consider the lagrangian of the program:

$$\begin{aligned} L &= F(q)\left[u(w_{H,1}) + \beta p_H u(w_{H,2})\right] + (1 - F(q))(u(z) + \beta p_H U_{1p}) \\ &\quad + \lambda \left\{H(q)\left[y - w_{H,1} + \beta p_H (y - w_{H,2})\right] - C\right\} \\ &\quad - \delta \left\{F(q)\left[u(w_{H,1}) + \beta p_L u(w_{H,2})\right] + (1 - F(q))(u(z) + \beta p_L U_{1p}) - U_L\right\} \end{aligned}$$

Solving the FOCs, with respect to w_1, w_2 and q:

$$\begin{aligned} F(q_H^{AI}) u'\left(w_{H,1}^{AI}\right) - \lambda H(q_H^{AI}) - \delta F(q_H^{AI}) u'\left(w_{H,1}^{AI}\right) &= 0 \\ F(q_H^{AI}) \beta p_H u'\left(w_{H,2}^{AI}\right) - \lambda \beta p_H H(q_H^{AI}) - \delta F(q_H^{AI}) \beta p_L u'\left(w_{H,2}^{AI}\right) &= 0 \\ F'(q_H^{AI}) \left[u\left(w_{H,1}^{AI}\right) + \beta p_H u\left(w_{H,2}^{AI}\right) - u(z) - \beta p_H U_{1p}\right] & \\ + \lambda H'(q_H^{AI}) \left[y - w_{H,1}^{AI} + \beta p_H \left(y - w_{H,2}^{AI}\right)\right] & \\ - \delta F'(q_H^{AI}) \left\{\left[u\left(w_{H,1}^{AI}\right) + \beta p_L u\left(w_{H,2}^{AI}\right) - (u(z) + \beta p_L U_{1p})\right]\right\} &= 0 \end{aligned}$$

We can rearrange the two first FOCs to obtain the following equations:

$$\begin{aligned} u'(w_{H,1}^{AI}) &= \frac{\lambda q}{1 - \delta} \\ u'(w_{H,2}^{AI}) &= \frac{\lambda q}{1 - \delta \frac{p_L}{p_H}} < \frac{\lambda q}{1 - \delta} = u'(w_{H,1}^{AI}) \end{aligned}$$

We deduce that *(i)* $w_{H,1}^{AI} < w_{H,2}^{AI}$ and *(ii)* $0 < \delta < 1$. The result that δ is greater than zero follows readily from the method of Lagrangian multiplier with inequality constraint.

Claim 2a: $w_{H,1}^{AI} < w_L^*$. Notice that since the allocation for the type L workers are not distorted, we have $w_L^* = w_{L,1}^{AI} = w_{L,2}^{AI}$, and $q_L^* = q_L^{AI}$. Now suppose $w_H^{AI} \geq w_L^*$. Because we already established $w_{H,1}^{AI} < w_{H,2}^{AI}$, then from the binding IC constraint

$$\begin{aligned} & F(q_H^{AI})\left[u(w_{H,1}^{AI}) + \beta p_L u(w_{H,2}^{AI})\right] + \left(1 - F(q_H^{AI})\right)(u(z) + \beta p_L U_{1p}) \\ =\ & F(q_L^{AI})\left[u(w_{L,1}^{AI}) + \beta p_L u(w_{L,2}^{AI})\right] + \left(1 - F(q_L^{AI})\right)(u(z) + \beta p_L U_{1p}) \\ =\ & U_L^{AI} \end{aligned}$$

We could conclude that $F\left(q_H^{AI}\right) \leq F\left(q_L^{AI}\right)$. Thus $q_H^{AI} \geq q_L^{AI}$.

The FOC with respect to q for type H could be rewritten in the following way:

$$
\begin{aligned}
&\frac{1}{1-\delta}\frac{1}{u'\left(w_{H,1}^{AI}\right)}\left\{\begin{array}{c}\left[u\left(w_{H,1}^{AI}\right)+\beta p_H u\left(w_{H,2}^{AI}\right)-u(z)-\beta p_H U_{1p}\right] \\ -\delta\left[u\left(w_{H,1}^{AI}\right)+\beta p_L u\left(w_{H,2}^{AI}\right)-u(z)-\beta p_L U_{1p}\right]\end{array}\right\} \\
= & -\frac{H'\left(q_H^{AI}\right)}{F'\left(q_H^{AI}\right)}\frac{1}{q_H^{AI}}\frac{C}{H\left(q_H^{AI}\right)}
\end{aligned}
$$

The FOC with respect to q for type L could be rewritten in the following way:

$$
\frac{1}{u'(w_L^*)}\left[u(w_L^*)+\beta p_L u(w_L^*)-u(z)-\beta p_L U_{1p}\right]=-\frac{H'(q_L^*)}{F'(q_L^*)}\frac{1}{q_L^*}\frac{C}{H(q_L^*)}
$$

First, we notice that, the RHS of both FOC is strictly decreasing in q. Because we can make use of the relationship $H(q)=F(q)q$ to rewrite it as

$$
-\frac{H'(q)}{F'(q)}\frac{1}{q}\frac{C}{H(q)}=-\frac{H'(q)}{\frac{H'(q)}{q}-\frac{H(q)}{q^2}}\frac{1}{q}\frac{C}{H(q)}=\underbrace{\frac{\varepsilon_H}{1-\varepsilon_H}}_{\text{decreasing in } q_H}\frac{C}{H(q)}
$$

Secondly, it could be noticed that if we substitute w_L on the LHS of the FOC for type H, it becomes:

$$
\begin{aligned}
LHS_{typeH}(w_L,w_L) &= \frac{1}{1-\delta}\frac{1}{u'(w_L^*)}\left\{\begin{array}{c}\left[u(w_L^*)+\beta p_H u(w_L^*)-u(z)-\beta p_H U_{1p}\right] \\ -\delta\left[u(w_L^*)+\beta p_L u(w_L^*)-u(z)-\beta p_L U_{1p}\right]\end{array}\right\} \\
&= \underbrace{\frac{1}{1-\delta}\frac{\beta(p_H-p_L)\left[u(w_L^*)-U_{1p}\right]}{u'(w_L^*)}}_{>0} \\
&\quad \underbrace{+\frac{1}{u'(w_L^*)}\left[u(w_L^*)+\beta p_L u(w_L^*)-u(z)-\beta p_L U_{1p}\right]}_{LHS_{typeL}}
\end{aligned}
$$

Thus $LHS_{typeH}(w_L^*,w_L^*)$ is always greater than the LHS of the FOC for type L, i.e. $LHS_{typeH}(w_L^*,w_L^*) > LHS_{typeL}(w_L^*,w_L^*)$. But if it is so, then $LHS_{typeH}\left(w_{H,1}^{AI},w_{H,2}^{AI}\right) > LHS_{typeH}(w_L^*,w_L^*)$, because we supposed $w_L^* \le w_{H,1}^{AI} < w_{H,2}^{AI}$.

These will imply that the RHS of the FOC for type H should be strictly greater than the RHS of the FOC for type L, i.e. $RHS_{typeH} > RHS_{typeL}$. Since the RHS of both FOCs have the same functional form, then it must be that $q_H^{AI} < q_L^*$, given that the RHS is decreasing in q. However, we established at the beginning of this proof that $q_H^{AI} \ge q_L^*$, we could reach the contradiction.

Claim 2b: $w_H^* < w_{H,2}^{AI}$. This could be shown using the same way as in the proof of claim 2a. Thus in total we have $w_{H,1}^{AI} < w_L^* < w_H^* < w_{H,2}^{AI}$.

Claim 3: $U_H^{AI} < U_H^*$. This is true because U_H^{AI} is the utility at the maximum level of an optimization program with an additional constraint, compared to U_H^*.

Claim 4: $q_H^{AI} < q_L^*$. We are going to proceed by contradiction. Suppose that $q_H^{AI} \geq q_L^*$. From the IC constraints

$$F\left(q_H^{AI}\right) \times \left(u\left(w_{H,1}^{AI}\right) + \beta p_L u\left(w_{H,2}^{AI}\right)\right) + \left(1 - F\left(q_H^{AI}\right)\right)\left(u(z) + \beta p_L U_{1p}\right) \leq U_L^*$$

we are able to obtain $u\left(w_{H,1}^{AI}\right) + \beta p_L u\left(w_{H,2}^{AI}\right) \geq (1 + \beta p_L)\, u(w_L^*)$.

Then we can use the same arugments as in the proof of proposition 4 in appendix D to show that, the first order constraints imply that $q_H^{AI} < q_L^*$. We hence obtain a contradiction. So we conclude that $q_H^{AI} < q_L^*$.

Claim 5: $q_H^* < q_H^{AI}$. Firstly, we have that in the equilibrium with asymmetric information, $w_{H,1}^{AI}$ is negatively correlated with q_H^{AI}, and $w_{H,2}^{AI}$ is positively correlated with q_H^{AI}.

Secondly, notice that when $p_L \to p_H$, the asymmetric information allocation should tend to the full information allocation by continuity. More precisely, since $w_{H,1}^{AI} < w_L^* < w_H^* < w_{H,2}^{AI}$, we know that $w_{H,1}^{AI}$ should gradually increase to w_H^* and $w_{H,2}^{AI}$ should gradually decrease to w_H^*.[26]

Now suppose that $q_H^{AI} \leq q_H^*$. So when $p_L \to p_H$, q_H^{AI} should increase to q_H^*. Then both $w_{H,1}^{AI}$ and q_H^{AI} increase with p_L, and it implies that $w_{H,1}^{AI}$ is positively correlated with q_H^{AI} - a contradiction.

On the other hand, if $q_H^* < q_H^{AI}$, then when $p_L \to p_H$, q_H^{AI} decreases to q_H^*. At the same time, $w_{H,2}^{AI}$ decreases to w_H^* and $w_{H,1}^{AI}$ increases to w_H^*. This is consistent with the fact that $w_{H,1}^{AI}$ is negatively correlated with q_H^{AI}, and $w_{H,2}^{AI}$ is positively correlated with q_H^{AI}.

Claim 6: Type H workers are worse off choosing type L's allocation, so that they do no have incentive to misrepresent themselves as type L. If type H workers choose the allocation of type L workers, the expected utility that they would enjoy is

$$F(q_L^*)\left(u(w_L^*) + \beta p_H u(w_L^*)\right) + \left(1 - F(q_L^*)\right)\left(u(z) + \beta p_H U_{1p}\right)$$

We want to show that this amount is smaller than the following which is the amount type H would secure if they stick to their own allocation:

[26](w_L^* and w_H^* will coincide when $p_L \to p_H$)

$$F\left(q_H^{AI}\right)\left(u\left(w_{H,1}^{AI}\right)+\beta p_H u\left(w_{H,2}^{AI}\right)\right)+\left(1-F\left(q_H^*\right)\right)\left(u(z)+\beta p_H U_{1p}\right)$$

The comparison could be simplified to the following

$$\begin{aligned} & F\left(q_L^*\right)\left[\left(u\left(w_L^*\right)+\beta p_H u\left(w_L^*\right)\right)-\left(u(z)+\beta p_H U_{1p}\right)\right] \\ < \; & F\left(q_H^{AI}\right)\left[\left(u\left(w_{H,1}^{AI}\right)+\beta p_H u\left(w_{H,2}^{AI}\right)\right)-\left(u(z)+\beta p_H U_{1p}\right)\right] \end{aligned}$$

According to the incentive constraint as follows

$$\begin{aligned} & F\left(q_L^*\right)\left[u\left(w_L^*\right)+\beta p_L u\left(w_L^*\right)-u(z)-\beta p_L U_{1p}\right] \\ = \; & F\left(q_H^{AI}\right)\left[u\left(w_{H,1}^{AI}\right)+\beta p_L u\left(w_{H,2}^{AI}\right)-u(z)-\beta p_L U_{1p}\right] \end{aligned}$$

, it suffices that the following inequality is true

$$\begin{aligned} & F\left(q_L^*\right)\left[u\left(w_L^*\right)-U_{1p}\right] \\ < \; & F\left(q_H^{AI}\right)\left[u\left(w_{H,2}^{AI}\right)-U_{1p}\right] \end{aligned}$$

On one hand, since $q_H^{AI} < q_L^*$, we have $F\left(q_H^{AI}\right) > F\left(q_L^*\right)$. On the other hand, since $w_{H,2}^{AI} > w_H^* > w_L^*$, we have $u\left(w_{H,2}^{AI}\right) > u\left(w_L^*\right)$. So it indeed holds true.

Claim 7: $U_H^{AI} > U_L^*$. It follows from the following relationships.

$$\begin{aligned} & U_L^* \\ = \; & F\left(q_H^{AI}\right)\left(u\left(w_{H,1}^{AI}\right)+\beta p_L u\left(w_{H,2}^{AI}\right)\right)+\left(1-F\left(q_H^*\right)\right)\left(u(z)+\beta p_L U_{1p}\right) \\ < \; & F\left(q_H^{AI}\right)\left(u\left(w_{H,1}^{AI}\right)+\beta p_H u\left(w_{H,2}^{AI}\right)\right)+\left(1-F\left(q_H^*\right)\right)\left(u(z)+\beta p_H U_{1p}\right) \end{aligned}$$

The equality comes from the incentive constraint, and the inequality comes from the fact that $p_L < p_H$.

1.6.5 Proof of Proposition 5

Consider a pooling allocation which treats both types identically. All firms post $w_{\text{1st period}} = w_{\text{2nd period}} = w_\alpha$. The queue length $q_\alpha = \alpha\frac{N}{M}+(1-\alpha)\frac{N}{M} = \frac{N_H}{M}+\frac{N_L}{M}$ will satisfy:

$$H\left(q_\alpha\right)\left[\left(y-w_\alpha\right)+\beta p_\alpha\left(y-w_\alpha\right)\right]=C$$

where $p_\alpha = \alpha p_H + (1-\alpha)p_L$. Furthermore, let (q_α, w_α) be the solution of the following program:

$$\begin{aligned} U_\alpha \; = \; & \max\left\{F(q)\left[u\left(w_\alpha\right)+\beta p_\alpha u\left(w_\alpha\right)\right]+(1-F(q))\left(u(z)+\beta p_\alpha U_{1p}\right)\right\} \\ \text{s.to} \quad & H(q)\left[y-w_\alpha+\beta p_\alpha\left(y-w_\alpha\right)\right]=C \end{aligned}$$

If we compare the above zero profit condition ($H(q_\alpha)[(y-w_\alpha)+\beta p_\alpha(y-w_\alpha)]=C$) to firms' zero profit condition in the separating equilibrium with asymmetric information, $H\left(q_H^{AI}\right)\left[\left(y-w_{1,H}^{AI}\right)+\beta p_H\left(y-w_{2,H}^{AI}\right)\right]=C$, we notice that, when $p_\alpha \to p_H$, i.e., when the fraction of bad workers is sufficiently low, the optimization program constructed just above approaches the optimization program for the type-H workers in the full information (FI) case where IC constraint is absent. This implies that, when $p_\alpha \to p_H$, we obtain $w_\alpha \to w_H^{FI}$ and $q_\alpha \to q_H^{FI}$, hence the good workers will be better off in terms of expected utility, when the fraction of bad workers is negligible

Furthermore, recall that in the full information case, the equilibrium allocation is such that $q_H^{FI} < q_L^{FI}$ and $w_H^{FI} > w_L^{FI}$. These two inequalities imply that the type L-workers could be employed more often, and receive higher ex post wages under this constructed pooling allocation. We could in turn conclude that the bad workers will also be better off in terms of expected utility under this pooling contract.

1.6.6 Proof of Proposition 6

Proof. the optimization program under full information is as follows. For firms which desire the type-H workers, we have

$$\begin{array}{rcl} U_H^*(h_H,z) & = & \max_{\{q,w_{H,1},w_{H,2}\}}\left\{\begin{array}{l} F(q_H)\left[u(w_{H,1})+\beta u(w_{H,2})\right] \\ +(1-F(q_H))\left[u(z)+\beta\max\left\{u(h_H),U_{1p}(h_H,z)\right\}\right] \end{array}\right\} \\ \text{s.t.} & & H(q_H)\left[y-w_{H,1}+\beta(y-w_{H,2})\right]=C \end{array}$$

where

$$\begin{array}{rcl} U_{1p}(h,z) & = & \max_{\{q,w\}}\left\{F(q)u(w)+(1-F(q))\left[u(h)+u(z)\right]\right\} \\ \text{s.t.} & & H(q)\left[y-w\right]=C \end{array}$$

and for the firms which desire the the type-L workers, we have

$$\begin{array}{rcl} U_L^*(h_L,z) & = & \max_{\{q,w_{L,1},w_{L,2}\}}\left\{F(q_L)u(w_L)+(1-F(q_L))u(z)+\beta u(h_L)\right\} \\ \text{s.t.} & & H(q_L)\left[y-w_L\right]=C \end{array}$$

We are going to proceed as follows. Firstly (Step 1), we show that when h_H tends to zero, the type-L workers will have incentive to mimic these type-H workers. Then (Step 2), we show that as h_H increases, the expected payoff from misrepresenting their types becomes less and less attractive for the type-L workers. Finally (Step 3), we show that for the type-H workers with h_H slightly below $\hat{h}$, the type-L workers do have incentive to mimic them; then by continuity, we are able to conclude that type-L workers are always better off by misrepresenting their types as any type-H workers.

Step 1. Consider the case where the value of h_H is small enough such that $h \to 0$. In that case, the type-H's problem coincides with the two-period workers' problem in the previous section. And the type-L's allocation coincides with the one-period workers' allocation. We have shown that the one-period workers will misrepresent their types when information is asymmetric. Step 2. We should show that as h_H increases towards $\hat{h}$, the extra payoff from choosing the allocation of these h_H types becomes less and less for those type-L workers. We shall firstly define $U_{h_H|h_L}$ as the expected payoff of a type-L worker who chooses the allocation of type-H worker. And we have

$$\begin{aligned} U_{h_H|h_L} &= F\left(q^*\left(h_H\right)\right)u\left(w^*\left(h_H\right)\right)+\left(1-F\left(q^*\left(h_H\right)\right)\right)u\left(z\right)+\beta u\left(h_L\right) \\ &\text{and} \\ U_{h_L|h_L} &= F\left(q_L^*\right)u\left(w_L^*\right)+\left(1-F\left(q_L^*\right)\right)u\left(z\right)+\beta u\left(h_L\right) \end{aligned}$$

Notice that in Step 1, we have when $h_H = 0$, $U_{h_H|h_L} > U_{h_L|h_L}$. We shall show that $U_{h_H|h_L}$ decreases with h_H, or equivalently we want to examine whether

$$\begin{aligned} \frac{\partial U_{h_H|h_L}}{\partial h_H} &= \left[\frac{\partial F(q_H^*)}{\partial q_H^*}\frac{dq_H^*}{dh_H}u\left(w_H^*\right)+F\left(q_H^*\right)u'\left(w_H^*\right)\frac{dw_H^*}{dh_H}\right]-\frac{\partial F(q_H^*)}{\partial q_H^*}\frac{dq_H^*}{dh_H}u\left(z\right) \\ &\underset{?}{<} 0 \end{aligned}$$

By Envelope theorem we have

$$\frac{\partial U_H^*}{\partial h_H} = \begin{cases} \left(1-F\left(q_H^*\right)\right)\times\beta u'\left(h_H\right) & \text{for } u\left(h_H\right)\geq U_{1p}\left(h_H,z\right) \\ \left(1-F\left(q_H^*\right)\right)\times\beta\left(1-F\left(q_{1p}\right)\right)u'\left(h_H\right) & \text{for } u\left(h_H\right)< U_{1p}\left(h_H,z\right) \end{cases}$$

This implies that by implicit differentiation of U_H^* with respect to h_H, we should have

$$\begin{aligned} 0 &= \frac{\partial F(q_H^*)}{\partial q_H^*}\frac{dq_H^*}{dh_H}\left[u\left(w_H^*\right)+\beta u\left(w_H^*\right)\right]+F\left(q_H^*\right)\left(1+\beta\right)u'\left(w_H^*\right)\frac{dw_H^*}{dh_H} \\ &\quad -\frac{\partial F(q_H^*)}{\partial q_H^*}\frac{dq_H^*}{dh_H}\left[u\left(z\right)+\beta\max\left\{u\left(h_H\right),U_{1p}\left(h_H,z\right)\right\}\right] \end{aligned}$$

We firstly rearrange the above expression in the following way to obtain

$$\begin{aligned} 0 = &\overbrace{\left[\frac{\partial F\left(q_H^*\right)}{\partial q_H^*}\frac{dq_H^*}{dh_H}u\left(w_H^*\right)+F\left(q_H^*\right)u'\left(w_H^*\right)\frac{dw_H^*}{dh_H}\right]}^{\text{Term A}} \\ &+\beta\times\left[\frac{\partial F(q_H^*)}{\partial q_H^*}\frac{dq_H^*}{dh_H}u\left(w_H^*\right)+F\left(q_H^*\right)u'\left(w_H^*\right)\frac{dw_H^*}{dh_H}\right] \\ &\underbrace{-\frac{\partial F\left(q_H^*\right)}{\partial q_H^*}\frac{dq_H^*}{dh_H}u\left(z\right)}_{\text{Term B}}+\beta\times\underbrace{\left(-\frac{\partial F\left(q_H^*\right)}{\partial q_H^*}\frac{dq_H^*}{dh_H}\max\left\{u\left(h_H\right),U_{1p}\left(h_H,z\right)\right\}\right)}_{\text{Term C}} \end{aligned}$$

Since it could be straightforwardly verified that both the equilibrium value of w_H^* increases and the equilibrium q_H^* increase with h_H (either by revealed preference or by implicit differentiation), we could conclude that Term A is negative, owing to the fact that Term B and Term C are both positive. We can then rearrange again the expression in the following way:

$$
\begin{aligned}
0 \;=\;& \overbrace{\left[\frac{\partial F\left(q_H^*\right)}{\partial q_H^*}\frac{dq_H^*}{dh_H}u\left(w_H^*\right)+F\left(q_H^*\right)u'\left(w_H^*\right)\frac{dw_H^*}{dh_H}\right]}^{\text{Term A}<0}\overbrace{-\frac{\partial F\left(q_H^*\right)}{\partial q_H^*}\frac{dq_H^*}{dh_H}u\left(z\right)}^{\text{Term B}>0} \\
&+\beta\left[\underbrace{\frac{\partial F\left(q_H^*\right)}{\partial q_H^*}\frac{dq_H^*}{dh_H}u\left(w_H^*\right)+F\left(q_H^*\right)u'\left(w_H^*\right)\frac{dw_H^*}{dh_H}}_{\text{Term A}<0}\underbrace{-\frac{\partial F\left(q_H^*\right)}{\partial q_H^*}\frac{dq_H^*}{dh_H}\max\left\{u\left(h_H\right),U_{1p}\left(h_H,z\right)\right\}}_{\text{Term C}>0}\right]
\end{aligned}
$$

Then since "Term A + Term B" is always smaller than "Term A + Term C", we could conclude that Term A + Term B <0 and Term A + Term C >0. So that we have proved that

$$
\begin{aligned}
\frac{\partial U_{h_H|h_L}}{\partial h_H} \;=\;& \overbrace{\left[\frac{\partial F\left(q_H^*\right)}{\partial q_H^*}\frac{dq_H^*}{dh_H}u\left(w_H^*\right)+F\left(q_H^*\right)u'\left(w_H^*\right)\frac{dw_H^*}{dh_H}\right]}^{\text{Term A}<0}\overbrace{-\frac{\partial F\left(q_H^*\right)}{\partial q_H^*}\frac{dq_H^*}{dh_H}u\left(z\right)}^{\text{Term B}>0} \\
<\;& 0
\end{aligned}
$$

which means that as h_H increases towards $\hat{h}$, the extra payoff from choosing the allocation of these h_H types becomes less and less for those type-L workers.

Step 3. At last, we would like to show that for values of h_H slightly below $\hat{h}$, the type L workers always have incentive to misrepresent their types. For values of h_H slightly below $\hat{h}^{FI}$, the expected payoff for the type-H workers is written as

$$
\begin{aligned}
U_H^*\left(h_H,z\right) &= \left\{\begin{array}{l} F\left(q_H^*\right)\left[u\left(w_H^*\right)+\beta u\left(w_H^*\right)\right] \\ +\left(1-F\left(q_H^*\right)\right)\left[u\left(z\right)+\beta u\left(h_H\right)\right] \end{array}\right\} \\
&= \left\{\begin{array}{l} F\left(q_H^*\right)u\left(w_H^*\right)+\left(1-F\left(q_H^*\right)\right)u\left(z\right) \\ +\beta F\left(q_H^*\right)\left[u\left(w_H^*\right)-u\left(h_H\right)\right]+\beta u\left(h_H\right) \end{array}\right\}
\end{aligned}
$$

We now compare this payoff with the payoff of the threshold type-L workers, which is

$$
U_L^*\left(\hat{h},z\right) = \left\{\begin{array}{l} F\left(q_L^*\right)u\left(w_L^*\right)+\left(1-F\left(q_L^*\right)\right)u\left(z\right) \\ +\beta u\left(\hat{h}\right) \end{array}\right\}
$$

Since $U_H^*(h_H) \lessapprox U_H^*(\hat{h}) = U_L^*(\hat{h})$, we then have

$$\begin{aligned} & F(q_H^*)u(w_H^*) + (1-F(q_H^*))u(z) + \beta F(q_H^*)[u(w_H^*) - u(h_H)] \\ \approx\ & F(q_L^*)u(w_L^*) + (1-F(q_L^*))u(z) \end{aligned}$$

and we can rearrange it to

$$\begin{aligned} & F(q_H^*)u(w_H^*) + (1-F(q_H^*))u(z) - [F(q_L^*)u(w_L^*) + (1-F(q_L^*))u(z)] \\ \approx\ & -\beta F(q_H^*)[u(w_H^*) - u(h_H)] \end{aligned}$$

Notice that the left-hand side of the quasi-equality represents the net gain from misrepresenting the type for a type-L workers, i.e., $\left[U_{h_H|h_L} - u(h_L)\right] - \left[U_{h_L|h_L} - u(h_L)\right]$, where $U_{h_H|h_L} = [F(q_H^*)u(w_H^*) + (1-F(q_H^*))u(z)]$ and $U_{h_L|h_L} = [F(q_L^*)u(w_L^*) + (1-F(q_L^*))u(z)]$. Now it suffices to show that on the right hand side, $[u(w_H^*) - u(h_H)] \approx \left[u(w_H^*) - u\left(\hat{h}^{FI}\right)\right]$ is smaller than zero. This is always true. Recall that we argued that an individual only chooses to enter into the one-period market whenever $u(h) - u(w_H)$ is strictly greater than zero, because there is a discrete drop of utility by changing from a two period wage contract $\left(w_{2-period}^*, q_{2-period}^*\right)$ to a one-period contract $\left(w_{1-period}^*, q_{1-period}^*\right)$ as we showed in the benchmark section - within a 2-period contract, the worker always gets matched more often and enjoys higher per-period wage.

1.6.7 Proof of Proposition 7

The first order conditions with respect to $w_{H,1}$, $w_{H,2}$ and q_H are:

$$\begin{aligned} F\left(q_H^{AI}\right)u'\left(w_{H,1}^{AI}\right) - \lambda H\left(q_H^{AI}\right) - \delta F\left(q_H^{AI}\right)u'\left(w_{H,1}^{AI}\right) &= 0 \\ F\left(q_H^{AI}\right)\beta u'\left(w_{H,2}^{AI}\right) - \lambda\beta H\left(q_H^{AI}\right) &= 0 \\ F'\left(q_H^{AI}\right)\left[u\left(w_{H,1}^{AI}\right) + \beta u\left(w_{H,2}^{AI}\right) - u(z) - \beta \max\left\{u(h_H), U_{1p}\right\}\right] & \\ +\lambda H'\left(q_H^{AI}\right)\left[y - w_{H,1}^{AI} + \beta\left(y - w_{H,2}^{AI}\right)\right] & \\ -\delta F'\left(q_H^{AI}\right)\left[u\left(w_{H,1}^{AI}\right) - u(z)\right] &= 0 \end{aligned}$$

where λ and δ are the Lagrangian multipliers associated with the first and the second constraint. Given that $H(q)/F(q) = q$, the FOC leads to

$$\begin{aligned} u'\left(w_{H,1}^{AI}\right) &= \frac{\lambda q_H^{AI}}{1-\delta} \\ u'\left(w_{H,1}^{AI}\right) &= \lambda q_H^{AI} \end{aligned}$$

implying that $w^{AI}_{H,1} < w^{AI}_{H,2}$. Suppose that $w^{AI}_{H,1} \geq w^{*}_{H}$, then since $w^{AI}_{H,1} < w^{AI}_{H,2}$, we have $w^{AI}_{H,2} > w^{AI}_{H,1} \geq w^{*}_{H}$. On the one hand, by the binding zero profit conditions

$$
\begin{aligned}
H(q^{*}_{H})((1+\beta)y - w^{*}_{H} - \beta w^{*}_{H}) &= 0 \\
H(q^{AI}_{H})((1+\beta)y - w^{AI}_{H,1} - \beta w^{AI}_{H,2}) &= 0
\end{aligned}
$$

we have $q^{AI}_{H} > q^{*}_{H}$. However, on the other hand, by the first order condition with respect to q, which could be rewritten in the following manner,

$$
\begin{aligned}
\frac{u'(w^{*}_{H}) - u(z)}{u'(w^{*}_{H})} + \beta \frac{u'(w^{*}_{H}) - \max\{u(h_H), U_{1p}\}}{u'(w^{*}_{H})} &= \frac{\varepsilon(q^{*}_{H})}{1-\varepsilon(q^{*}_{H})} \frac{C}{H(q^{*}_{H})} \\
\frac{u'(w^{AI}_{H,1}) - u(z)}{u'(w^{AI}_{H,1})} + \beta \frac{u'(w^{AI}_{H,2}) - \max\{u(h_H), U_{1p}\}}{u'(w^{AI}_{H,2})} &= \frac{\varepsilon(q^{AI}_{H})}{1-\varepsilon(q^{AI}_{H})} \frac{C}{H(q^{AI}_{H})}
\end{aligned}
$$

we then could conclude that $q^{AI}_{H} < q^{*}_{H}$, because $\Phi(w) \equiv \frac{u(w)-u(z)}{u'(w)}$ is increasing in w and $\Psi(q) \equiv \frac{\varepsilon(q)}{1-\varepsilon(q)} \frac{C}{H(q)}$ is decreasing in q. We then reach a contradiction, so that $w^{AI}_{H,1} < w^{*}_{H}$. To prove $w^{*}_{H} < w^{AI}_{H,2}$, we could proceed by exactly the same argument.

The other results can be similarly reproduced as in Proposition 4, given that this program is just a particular case of the problem there with $p_H = 1$ and $p_L = 0$.

1.6.8 Proof of Lemma 1

(1) Incentive constraint implies that

$$
\begin{aligned}
e(h)u(w_2(h)) - e(h)u(h) &\geq e(\tilde{h})u(w_2(\tilde{h})) - e(\tilde{h})u(h) \\
&\text{and} \\
e(\tilde{h})u(w_2(\tilde{h})) - e(\tilde{h})u(\tilde{h}) &\geq e(h)u(w_2(h)) - e(h)u(\tilde{h})
\end{aligned}
$$

Adding these two together, we obtain

$$
\left(e(h) - e(\tilde{h})\right)\left(u(\tilde{h}) - u(h)\right) \geq 0
$$

Hence $e(h)$ is non-increasing in h. And it is also differentiable almost everywhere, from which $w_2(h)$ is also differentiable with the same points of non-differentiability.

(2) By standard argument as in Laffont and Martimort (2001), it could be verified that the local incentive constraints imply the global incentive constraints. So that $J(h,h) = \max_{\tilde{h}} J(h,\tilde{h})$. This implies that we can use the Envelope condition to obtain

$$
\frac{\partial J(h,h)}{\partial h} = -e(h)u'(h) \leq 0
$$

By integration with respect to h, we obtain

$$J(h,h) - J\left(\overline{h},\overline{h}\right) \quad = \quad \int_h^{\overline{h}} e(x)\, u^{'}(x)\, dx$$

(3) Since $\frac{\partial J(h,h)}{\partial h} \leq 0$, it suffices to impose $J\left(\overline{h},\overline{h}\right) \geq 0$ to make sure that all types participate.

by virtue of the fact that $0 < F(\hat{q}) < 1$. We hence obtain the following relationship:

1.6.9 Proof of Proposition 9

We have proved above that $w_1^{AI} < w_2^{AI} < y = w^{FI}$. Now we prove $\hat{h}^{AI} < \hat{h}^{FI}$. Suppose $\hat{h}^{AI} \geq \hat{h}^{FI}$. Then it implies that $\hat{h}^{AI} > w_2^{AI}$. Then, according to the optimal solution of firms' retaining choice, there exists some value of h with $\hat{h}^{AI} > h > w_2^{AI}$ such that $e(h) = 1$. However, this contradicts the individual rationality constraint which requires that $J(h,h) = e(h) \times \left(u\left(w_2^{AI}\right) - u(h)\right) \geq 0$ for all h.

Now we prove $q^{AI} > q^{FI}$. Firstly, we have the following FOC with respect to q^{AI} in case of asymmetric information:

$$\begin{aligned}\frac{u\left(w_1^{AI}\right)-u(z)}{u^{'}\left(w_1^{AI}\right)} + \beta \frac{\left\{G\left(\hat{h}^{AI}\right)u\left(w_2^{AI}\right)+\int_{\hat{h}^{AI}}^{\overline{h}} u(h)dG(h)-\int_{\underline{h}}^{\overline{h}} \max\left\{U_{1p},u(h)\right\}dG(h)\right\}}{u^{'}\left(w_1^{AI}\right)} & \\ y - w_1^{AI} + \beta G\left(\hat{h}^{AI}\right)\left(y - w_2^{AI}\right) &= \frac{1}{1-\varepsilon\left(q^{AI}\right)} \frac{C}{H\left(q^{AI}\right)}\end{aligned}$$

Secondly, we define that q^{FI} as being solved by the following program:

$$\begin{aligned} U^{FI} \quad = \quad & F(q)\left\{u(y) + \beta \int_{\underline{h}}^{y} \left[u(y) - u(h)\right] dG(h) + \beta \int_{\underline{h}}^{\overline{h}} u(h)\, dG(h)\right\} \\ & + (1 - F(q))\left\{u(z) + \beta \int_{\underline{h}}^{\overline{h}} \max\left\{U_{1p}, u(h)\right\} dG(h)\right\} - \frac{1}{q} C \times u^{'}(y)\end{aligned}$$

And the FOC with respect to q gives

$$\begin{aligned}\frac{u(y)-u(z)}{u^{'}(y)} + \beta \frac{\left\{G(y)u(y)+\int_y^{\overline{h}} u(h)dG(h)+\int_{\underline{h}}^{\overline{h}}\left[u(h)-\max\left\{U_{1p},u(h)\right\}\right]dG(h)\right\}}{u^{'}(y)} & \\ +y - y + \beta G(y)(y - y) &= \frac{1}{1-\varepsilon\left(q^{FI}\right)} \frac{C}{H\left(q^{FI}\right)}\end{aligned}$$

We observe three elements:

1. the "LHS" of the FOC for the Asymmetric-information economy is increasing in w;

2. when $w_1 = w_2 = y$, the "LHS" of the FOC for the Asymmetric-information economy coincides with the "LHS" of FOC for the Full-information economy;

3. the "RHS" of both equations are both of functional form and decreasing in q.

Thus to show $q^{FI} < q^{AI}$, it is sufficient to show that $LHS^{FI} > LHS^{AI}$. We will proceed in the following way:

(i) Since $w_1^{AI} < y$, we could obtain $\frac{u\left(w_1^{AI}\right)-u(z)}{u'\left(w_1^{AI}\right)} + y - w_1^{AI} < \frac{u(y)-u(z)}{u'(y)} + y - y$, because $\frac{u\left(w_1^{AI}\right)-u(z)}{u'\left(w_1^{AI}\right)} + y - w_1^{AI}$ is increasing in w_1^{AI}.

(ii) In the following, The first inequality comes from the facts that $w_1^{AI} < w_2^{AI}$ and $u'(w)$ being decreasing, and the second inequality is due to the same argument stated in *(i)*:

$$\begin{aligned}
& \frac{\left\{G\left(\hat{h}^{AI}\right)u\left(w_2^{AI}\right)+\int_{\hat{h}^{AI}}^{\bar{h}} u(h)dG(h)-\int_{\underline{h}}^{\bar{h}} \max\left\{U_{1p},u(h)\right\}dG(h)\right\}}{u'\left(w_1^{AI}\right)} + G\left(\hat{h}^{AI}\right)\left(y-w_2^{AI}\right) \\
< \; & \frac{\left\{G\left(\hat{h}^{AI}\right)u\left(w_2^{AI}\right)+\int_{\hat{h}^{AI}}^{\bar{h}} u(h)dG(h)-\int_{\underline{h}}^{\bar{h}} \max\left\{U_{1p},u(h)\right\}dG(h)\right\}}{u'\left(w_2^{AI}\right)} + G\left(\hat{h}^{AI}\right)\left(y-w_2^{AI}\right) \\
< \; & \frac{\left\{G(y)u(y)+\int_{\hat{h}^{AI}}^{\bar{h}} u(h)dG(h)-\int_{\underline{h}}^{\bar{h}} \max\left\{U_{1p},u(h)\right\}dG(h)\right\}}{u'(y)} + G(y)(y-y)
\end{aligned}$$

Thus the $LHS^{FI} > LHS^{AI}$, which implies $RHS^{FI} > RHS^{AI}$, and it turns out that $q^{FI} < q^{AI}$ by the fact that the right hand sides of the above FOCs are of the same functional form and decreasing in q.

To conclude, it implies that $F(q^{AI})G(\hat{h}^{AI}) < F(q^{FI})G(\hat{h}^{FI})$, the job matching is inferior in the economy with asymmetric information.

1.6.10 Proof of Proposition 1.6.10

We use δ as the Lagrangian multiplier for the incentive constraint. The first order condition with respect to $w_{L,1}$, $w_{L,2}$ and $w_{L,3}$ gives respectively:

$$\begin{aligned}
F\left(q_l^{AI}(H)\right)u'\left(w_{l,1}^{AI}(H)\right) - \lambda_l^{AI}(H)H\left(q_l^{AI}(H)\right) - \delta_l^{AI}(H)F\left(q_l^{AI}(H)\right)u'\left(w_{l,1}^{AI}(H)\right) &= 0 \\
F\left(q_l^{AI}(H)\right)\beta p_H u'\left(w_{l,2}^{AI}(H)\right) - \lambda_l^{AI}(H)\beta p_H H\left(q_l^{AI}(H)\right) \\
-\delta_l^{AI}(H)F\left(q_l^{AI}(H)\right)\beta p_L u'\left(w_{l,2}^{AI}(H)\right) &= 0 \\
F\left(q_l^{AI}(H)\right)\beta^2 p_H^2 u'\left(w_{l,3}^{AI}(H)\right) - \lambda_l^{AI}(H)\beta^2 p_H^2 H\left(q_l^{AI}(H)\right) \\
-\delta_l^{AI}(H)F\left(q_l^{AI}(H)\right)\beta^2 p_L^2 u'\left(w_{l,3}^{AI}(H)\right) &= 0
\end{aligned}$$

Rearranging, we obtain

$$
\begin{aligned}
u^{'}\left(w_{l,1}^{AI}(H)\right) &= \frac{\lambda_l^{AI}(H)q_l^{AI}(H)}{1-\delta_l^{AI}(H)}\\
u^{'}\left(w_{l,2}^{AI}(H)\right) &= \frac{\lambda_l^{AI}(H)q_l^{AI}(H)}{1-\frac{p_L}{p_H}\delta_l^{AI}(H)}\\
u^{'}\left(w_{l,3}^{AI}(H)\right) &= \frac{\lambda_l^{AI}(H)q_l^{AI}(H)}{1-\frac{p_L}{p_H}\frac{p_L}{p_H}\delta_l^{AI}(H)}
\end{aligned}
$$

Since $\frac{p_L}{p_H}\frac{p_L}{p_H} < \frac{p_L}{p_H} < 1$, we could conclude that we have $w_{l,1}^{AI}(H) < w_{l,2}^{AI}(H) < w_{l,3}^{AI}(H)$. Notice that now we have three points of wages, instead of merely two points as in Section 2. A natural question to ask is how wage grows. We first prove the following claim, then show that when $u^{'''} < 0$, the wage growth is concave.

Claim. $u^{'}\left(w_{l,1}^{AI}(H)\right) - u^{'}\left(w_{l,2}^{AI}(H)\right) > u^{'}\left(w_{l,2}^{AI}(H)\right) - u^{'}\left(w_{l,3}^{AI}(H)\right)$.

Proof of Claim. From the above first order condition, we have

$$
\begin{aligned}
\frac{u^{'}\left(w_{l,1}^{AI}(H)\right)}{u^{'}\left(w_{l,2}^{AI}(H)\right)} &= \frac{1-\frac{p_L}{p_H}\delta_l^{AI}(H)}{1-\delta_l^{AI}(H)}\\
\frac{u^{'}\left(w_{l,2}^{AI}(H)\right)}{u^{'}\left(w_{l,3}^{AI}(H)\right)} &= \frac{1-\frac{p_L}{p_H}\frac{p_L}{p_H}\delta_l^{AI}(H)}{1-\frac{p_L}{p_H}\delta_l^{AI}(H)}
\end{aligned}
$$

Denote $\frac{p(L)}{p(H)} = \kappa < 1$, then

$$
\frac{u^{'}\left(w_{l,1}^{AI}(H)\right)}{u^{'}\left(w_{l,2}^{AI}(H)\right)} - \frac{u^{'}\left(w_{l,2}^{AI}(H)\right)}{u^{'}\left(w_{l,3}^{AI}(H)\right)} = \frac{\left(\kappa^2+1-2\kappa\right)\delta}{(1-\kappa\delta)(1-\delta)} > 0 \quad \Rightarrow \quad \frac{u^{'}\left(w_{l,1}^{AI}(H)\right)}{u^{'}\left(w_{l,2}^{AI}(H)\right)} > \frac{u^{'}\left(w_{l,2}^{AI}(H)\right)}{u^{'}\left(w_{l,3}^{AI}(H)\right)}
$$

We subtract 1 at both sides, and we obtain

$$
\frac{u^{'}\left(w_{l,1}^{AI}(H)\right) - u^{'}\left(w_{l,2}^{AI}(H)\right)}{u^{'}\left(w_{l,2}^{AI}(H)\right)} > \frac{u^{'}\left(w_{l,2}^{AI}(H)\right) - u^{'}\left(w_{l,3}^{AI}(H)\right)}{u^{'}\left(w_{l,3}^{AI}(H)\right)}
$$

Multiply $u^{'}\left(w_{l,2}^{AI}(H)\right)$ at both sides. Since $w_{l,2}^{AI}(H) < w_{l,3}^{AI}(H)$, and $u^{'}$ is decreasing in w, we obtain

$$
u^{'}\left(w_{l,1}^{AI}(H)\right) - u^{'}\left(w_{l,2}^{AI}(H)\right) > u^{'}\left(w_{l,2}^{AI}(H)\right) - u^{'}\left(w_{l,3}^{AI}(H)\right)
$$

Now we show that the wage growth is concave under $u^{'''}(.) < 0$. When $u^{'''}(.) < 0$, $u^{'}(.)$ is concave. By the property of concave functions, we have

$$u^{'}\left(w_{l,1}^{AI}(H)\right) - u^{'}\left(w_{l,2}^{AI}(H)\right) < \left(w_{l,1}^{AI}(H) - w_{l,2}^{AI}(H)\right) \times u^{''}\left(w_{l,2}^{AI}(H)\right)$$

Similarly, we have

$$\begin{aligned} u^{'}\left(w_{l,3}^{AI}(H)\right) - u^{'}\left(w_{l,2}^{AI}(H)\right) &< \left(w_{l,3}^{AI}(H) - w_{l,2}^{AI}(H)\right) \times u^{''}\left(w_{l,2}^{AI}(H)\right) \\ &\Updownarrow \\ u^{'}\left(w_{l,2}^{AI}(H)\right) - u^{'}\left(w_{l,3}^{AI}(H)\right) &> \left(w_{l,2}^{AI}(H) - w_{l,3}^{AI}(H)\right) \times u^{''}\left(w_{l,2}^{AI}(H)\right) \end{aligned}$$

Hence, we have

$$\begin{aligned} & \left(w_{l,1}^{AI}(H) - w_{l,2}^{AI}(H)\right) \times u^{''}\left(w_{l,2}^{AI}(H)\right) \\ > \quad & u^{'}\left(w_{l,1}^{AI}(H)\right) - u^{'}\left(w_{l,2}^{AI}(H)\right) \\ > \quad & u^{'}\left(w_{l,2}^{AI}(H)\right) - u^{'}\left(w_{l,3}^{AI}(H)\right) \\ > \quad & \left(w_{l,2}^{AI}(H) - w_{l,3}^{AI}(H)\right) \times u^{''}\left(w_{l,2}^{AI}(H)\right) \end{aligned}$$

Then we divide $-u^{''}\left(w_{l,2}^{AI}(H)\right)$ at both sides of $\begin{aligned} & \left(w_{l,1}^{AI}(H) - w_{l,2}^{AI}(H)\right) \times u^{''}\left(w_{l,2}^{AI}(H)\right) \\ > & \left(w_{l,2}^{AI}(H) - w_{l,3}^{AI}(H)\right) \times u^{''}\left(w_{l,2}^{AI}(H)\right) \end{aligned}$, we could obtain $w_{l,2}^{AI}(H) - w_{l,1}^{AI}(H) > w_{l,3}^{AI}(H) - w_{l,2}^{AI}(H)$.

Chapter 2

Age-directed Policy in Competitive Search Equilibrium under Asymmetric Information

Coauthored with Professor François Langot

Abstract

Based on Bi and Langot (2015), we study the welfare impact of redistributively front-loading unemployment benefit by age. In the two-period context, such a transfer relaxes information constraint, leading to welfare improvement of the low type and aggregate welfare improvement. There may be, however, welfare loss for the high type when their cross-subsidization towards the low type is considerable enough. We find also that such a transfer can downwards distort the participation margin if workers make participation choice before search, leading to aggregate output loss. In the three-period benchmark, our result could be generalized under a fairly strong condition.

Keywords: Competitive search equilibrium, Asymmetric information, Age-targeted Policies

JEL classification: D82, D86, J14, J26, J64

2.1 Introduction

It is usual to present the optimal unemployment insurance system as an arbitration between "smoothing" and "incentive": the benefits allow risk averse workers to smooth their consumption flows, but the timing of these benefits must also lead them to accept quickly a job proposal (see Shavell and Weiss (1979)). In a principal-agent context with moral hazard, Hopenhayn and Nicolini (1997) (hereafter HN) show that the optimal contract consists of a regressive unemployment benefit which decreases with unemployment spell, as well as a wage tax which depends positively on the previous unemployment spell after reemployment. Hairault, Langot, Menard and Sopraseuth (2012) (hereafter HLMS) show that HN's result is weakened if age heterogeneity is present. Indeed, with an absorbing state retirement introduced, unemployment workers who are close to retirement have high probability of leaving permanently the labor force which prevents the principal from recouping the unemployment benefit transferred to these workers. A pension tax is suggested to remedy the problem. Michelacci and Ruffo (2015) pursue further through the direction of age-dependent policy. They show that redistributing unemployment benefit from the older to the younger workers can bring further gain to the welfare, even if the existing level of unemployment benefits is already close to optimal. Their result is based on two arguments. Firstly, unemployment insurance is relatively more valuable for the young workers, because consumption smoothing is harder to achieve for the young during a spell of unemployment; Secondly, the issue of moral hazard for the young generation is relatively mild, because the young should have stronger incentive of improving their career prospects - their human capital investments have higher marginal returns. The aim of this paper is to complete the analysis of an unemployment benefit system based on the worker age, by proposing an argument based on an equilibrium view of the labor market. Our main argument is the following: by providing additional revenues to unemployed young workers, the insurance system can better prevent the "bad" workers (the workers that will quit sooner) from applying to "good" jobs (the jobs that expect a longer employment relationship) in a market where firms lack information concerning the workers' horizon in the labor market (private information). Thus, a redistribution of the unemployment benefits from young to old workers could reduce the information costs.

Our equilibrium search framework is closest to Guerrieri (2008) and Guerrieri, Shimer and Wright (2010), and can be summarized as follow. There is only one source of heterogeneity: the expected horizon of remaining in the firm. Hiring a worker with longer horizon generates higher expected surplus, hence firms which aim at entering into such long term relationship are ready to sink more virtual cost ex ante by waiting longer for the match with such a worker to take place. However, the workers with short horizon are better off misrepresenting themselves as workers with longer horizon. The employers are then obliged to post contracts which induce self-selection - an adverse selection problem then arises. Since there is no means of further increasing the utility of workers with short horizon to make them less willing to practice arbitrage, the workers with long horizon's allocation must be distorted for this sake. Intuitively, these workers with long horizon generate higher expected surplus, so it is possible for the firm to sacrifice part of it for the screening process. Hence, a sorting mechanism is implemented: Bi and Langot (2015) prove the existence of segmented labor market equilibria when information is asymmetric. The incentive compatible contracts can endogenously segment the labor market: the "bad" (workers with short horizon) receive non-distorted contracts, whereas the "good" (those with longer horizon) receive distorted contracts. Although the discrimination process is directed "against" the bad workers, the induced cost is "paid for" by the good workers through a distorted contract. The bad workers receive information rent. The precise structure and context of Bi and Langot (2015) are as follows. In the benchmark, the workers are modelled as having an exogenous probability of quitting the labor market. This probability can be high or low, and is perfectly known at the time of hiring only by the workers.[1] We then extend the basic setup to two further frameworks where the workers are allowed to make participation choice. In the first framework with participation choice, workers receive, before search, a precise but private signal on their value of home production for their later stage of life, according to which they choose the submarket characterized by the length of employment relationship based on this private information.[2] In the second, the heterogeneity is only realised ex post - neither the workers nor the firms know the worker's utility of leisure before search, but during the contract this characteristic becomes observable by the worker.[3] At last, we consider a three-period extension of the benchmark (without participation choice), where the participation shock takes place recurrently and workers who are once shocked out of the market can become active again, to check the robustness of the model.

[1]This first case is close to many models of labor market where the agents' horizons are bounded: see Cheron, Hairault and Langot (2011, 2013), or Menzio, Telyukova and Visschers (2015).

[2]Since the middle of the 90s, the retirement choices are central for the policy maker, given the problems associated to financing of the Pay-As-You-Go system. Theoretical and empirical works also show how the labor market status interacts with this retirement choice: see Coile and Levine (2006, 2007, 2011), Hairault, Langot and Sopraseuth (2010) and Hairault, Langot and Zylberberg (2015).

[3]This model accounts for information that becomes available after a first part of career: workers have heterogeneous working conditions, leading to a heterogeneity in the wealth and/or health distributions, and thus, in the value of leisure in the second part of their careers. These worker-specific experiences may explain heterogeneity in the age of entry in disability programs (see e.g. Low and Pistaferri (2010)), or in the retirement age among ex ante identical individuals (see e.g. Cutler, Meara and Richards-Shubik (2011)).

In this paper, our focus is on the impact of redistribution through front-loading unemployment benefit by age. Since the policy maker is subject to the same asymmetric information problem as in the decentralized case, it necessarily involves cross-subsidization across different types of workers. Qualitatively, redistribution induces cross-subsidization of welfare across workers of different types, which can aggravate (or mitigate) the informational externality - misrepresenting one's type may become more (or less) attractive for the low type. Quantitatively, the extent of such cross-subsidization will depend upon the fraction of each type of workers who enjoy the unemployment benefit (hereafter UB) in the labor market. Some results are summarized as follows. In the two-period benchmark, where the shock occurs only once, redistribution is always beneficial for the "bad"; as for the "good", their welfare variation results from a comparison between the loss from cross-subsidization and a gain from mitigated information externality, and the latter effect dominates when the fraction of the "good" is sufficiently high. The aggregate welfare always improves, because the cross subsidization always makes the "bad" comparatively less reluctant to misrepresent their type - the information constraint is relaxed. Now we allow submarket choice. If this choice is made before search, the results from the benchmark on the welfare are not altered, however, we find that the participation margin is further distorted after redistribution in the sense that redistribution induces larger fraction of workers to choose the one-period submarket (so that they do not work in the second period). We argue that this could in turn reduce the aggregate output in the two-period submarket. Hence in this framework, there is potential trade-off between loss in the aggregate output and gain in aggregate welfare. Now consider that the participation choice is made after search. When unemployment benefit can not be claimed by the workers who quit prematurely to enjoy home production, such a transfer has no impact on the welfare since it does not enter into the information constraint. When it can be claimed, such redistribution can bring aggregate welfare gain, and it can increase the participation margin in case of asymmetric information. At last, from the study of a three-period extension of the benchmark with recurrent participation shock, we establish that the results from the two-period model could be generalized under a fairly strong condition.

The paper is organized as follows. In Section 1, we present the benchmark with exogenous participation choices. In Section 2, we introduce workers' participation choice which depends on their home production shock before search. We move to the situation where such a shock occurs only after matching, this is done in Section 3. Section 4 discusses the robustness, and then we conclude.

2.2 Exogenous quitting probability and age-specific UB

2.2.1 Model Assumptions

Consider a two-period economy with search friction. The number of workers is fixed, whereas the number of firms is variable and to be determined in the equilibrium. All workers start unemployed, and get hired by a firm through a job-search process. Denote q the tightness of the labor market, i.e. the worker/firm ratio. The hiring probability for the firms is $H(q)$, while $F(q)$ represents workers' job finding rate. As is usual in the matching models, we assume a constant return to scale relationship implying that *(i)* $H(q) = qF(q)$, *(ii)* the workers' job finding rate $F(q)$ is decreasing and concave in q, *(iii)* the firms' hiring probability $H(q)$ is increasing and concave in q, and *(iv)* the elasticity of firms' hiring probability is a decreasing function in q and between $(0,1)$, i.e. $\varepsilon'(q) = \left(\frac{H'(q)}{H(q)/q}\right)' < 0$ and $\varepsilon(q) \in (0,1)$.

We assume that an employed worker at the second period may leave the firm with certain probability for exogenous reasons, hence terminating the employment relationship prematurely. We focus on two possible types of worker: a fraction α will stay with probability p_H - type "H"; while the rest are type "L" workers, who stay with probability p_L, with $p_H > p_L$.

Such a wage posting game has three stages: *(i)* firms announce wages; *(ii)* workers observe the posted wages, and choose where to apply; *(iii)* the workers who get matched with firms produce, and payoffs are realized. With full information, Bi and Langot (2015) show that a competitive search equilibrium is characterized by *(i)* an optimal level of per period wage such that $w_H^{FI} > w_L^{FI}$, *(ii)* an optimal level of expected queue length such that $q_H^{FI} > q_L^{FI}$, *(iii)* the wages increase less faster than the increase of productivity with respect to p: $(1+\beta p_H)w_H^{FI} - (1+\beta p_L)w_L^{FI} < (1+\beta p_H)y - (1+\beta p_L)y$. Moreover, the workers who expect to stay longer in the firm enjoy higher expected utility from job search, even if different types offer the same quality of labor. This situation generates incentives for the type L workers to visit the submarket for the type H, if information is not perfectly shared. Thus firms which post long-term employment contracts and are uninformed about the workers' types may be unable to recoup their vacancy cost if they aim at attracting type H but end up with type L worker. This leads us to consider the possibility for firms to manipulate wage profile to induce self-selection among the workers, such that no workers will misrepresent their types.

2.2.2 Equilibrium

The optimization program for firms that want to attract p_H-type workers is written as follow:

$$\begin{array}{lcl} U_H^{AI} & = & \max_{q,w_{H,1},w_{H,2}} \left\{F(q)\left[u(w_{H,1})+\beta p_H u(w_{H,2})\right]+(1-F(q))\left[u(z)+\beta p_H U_{1p}\right]\right\} \\ \text{s.t.} & & H(q)\left[y-w_{H,1}+\beta p_H (y-w_{H,2})\right]=C \\ IC \text{ constraint} & & F(q_H)\left[u(w_{H,1})+\beta p_L u(w_{H,2})\right]+(1-F(q_H))\left[u(z)+\beta p_L U_{1p}\right] \le U_L^* \end{array}$$

where the optimal value of U_L is deduced from the optimization program for firms that want to attract type L workers. This problem is written as:

$$\begin{array}{lcl} U_L^* & = & \max_{q,w_{L,1},w_{L,2}} \left\{F(q)\left[u(w_{L,1})+\beta p_L u(w_{L,2})\right]+(1-F(q))\left[u(z)+\beta p_L U_{1p}\right]\right\} \\ \text{s.t.} & & H(q)\left[y-w_{L,1}+\beta p_L (y-w_{L,2})\right]=C \end{array}$$

Finally, for the one-period workers, the optimal contract is deduced from:

$$\begin{array}{lcl} U_{1p} & = & \max_{q,w} \left\{F(q)u(w)+(1-F(q))u(z)\right\} \\ \text{s.t.} & & H(q)\left[y-w\right]=C \end{array}$$

This economy is an extension of the theory proposed by Guerrieri, Shimer and Wright (2010). We will use the superscript "*AI*" for the equilibrium allocations of the program with incentive constraint, and the superscript "$*$" for the equilibrium allocations of the program without information constraint. The properties of this equilibrium are the following. Firstly, the separating equilibrium always exists. Secondly, the submarket allocation for type-L workers is not distorted, implying that $q_L^* = q_L^{FI}$, $w_{L,1}^* = w_{L,2}^* = w_L^{FI}$, and $U_L^* = U_L^{FI} < U_H^{AI}$, where FI stands for full information. Thirdly, the distortions are supported by the type H workers, leading to $w_{H,1}^{AI} < w_L^{FI} < w_H^{FI} < w_{H,2}^{AI}$, $q_H^{FI} < q_H^{AI} < q_L^{FI}$ and $U_L^{FI} < U_H^{AI} < U_H^{FI}$.

2.2.3 Redistributive age-specific UB

Now, we distinguish the first-period unemployment benefit z_1 and the second-period unemployment benefit z_2. The policy maker will redistribute by front-loading the unemployment benefit from the old to young. Starting with $z_1 = z_2 = z$. we lower the second period z_2 such that it becomes equivalent to z minus dz, i.e. $z_2 = z - dz$.

The government will expect to gather money from the following fraction of the old generation:

$$\alpha \times \left[\beta p_H \left(1-F\left(q_H^{AI}\right)\right)\left(1-F\left(q_{1p}^*\right)\right)\right]+(1-\alpha)\times\left[\beta p_L \left(1-F\left(q_L^*\right)\right)\left(1-F\left(q_{1p}^*\right)\right)\right] \tag{2.1}$$

where α is the fraction of type H workers. At the same time, the government will expect to redistribute to the following fraction of young generation who suffers from unemployment:

$$\alpha \times \left(1 - F\left(q_H^{AI}\right)\right) + (1-\alpha) \times \left(1 - F\left(q_L^*\right)\right) \tag{2.2}$$

To understand these expressions, notice that the expected fraction of workers who receive the first period $u(z_1)$ consists of those who are not able to find a job in the first period, and this fraction consists of $\left(1 - F\left(q_H^{AI}\right)\right)$ for type H, and $\left(1 - F\left(q_L^*\right)\right)$ for type L. For similar reason, the expected fraction of workers who receive the second period $u(z_2)$ consists of $\beta p_H \left(1 - F\left(q_H^{AI}\right)\right)\left(1 - F\left(q_{1p}^*\right)\right)$ for type H, and $\beta p_L \left(1 - F\left(q_L^*\right)\right)\left(1 - F\left(q_{1p}^*\right)\right)$ for type L. To keep budget balanced, every unit of decrease of old generation's unemployment benefit dz should be translated to

$$k \equiv \frac{\alpha \times \left[\beta p_H \left(1-F\left(q_H^{AI}\right)\right)\left(1-F\left(q_{1p}^*\right)\right)\right] + (1-\alpha) \times \left[\beta p_L \left(1-F\left(q_L^*\right)\right)\left(1-F\left(q_{1p}^*\right)\right)\right]}{\alpha \times \left(1-F\left(q_H^{AI}\right)\right) + (1-\alpha) \times \left(1-F\left(q_L^*\right)\right)} \times dz$$

unit of utility gain for the young generation regardless of the type. Such a transfer is justified by the fact that the planner has also no information on the agents' types.[4] Notice that it could be straightforwardly verified that we have $\beta p_L < \frac{k}{\left(1-F\left(q_{1p}^*\right)\right)} < \beta p_H$.

Proposition 1. *In this case of ex ante heterogeneity on quitting probability, redistributing unemployment benefit from the old to the young workers will have the following impacts:*

(1) type L workers' welfare increases; (2) type H workers' welfare decreases when α is sufficiently low, and increases when α is sufficiently high; (3) the aggregate welfare always increases.

Proof. (1). By Envelope Theorem, the increase of first-period z brings type L welfare gain by an amount of $\triangle_{L,gain}^{AI} = \left(1 - F\left(q_L^*\right)\right) \times u'(z) \times kdz$. At the same time, the decrease of second-period z brings type L welfare loss by an amount of $\triangle_{L,loss}^{AI} = \beta p_L \left(1 - F\left(q_L^*\right)\right)\left(1 - F\left(q_{1p}^*\right)\right) \times u'(z) \times dz$. So that

$$\begin{aligned} & \text{Sign}\left\{\triangle_{L,gain}^{AI} - \triangle_{L,loss}^{AI}\right\} \\ = \ & \text{Sign}\left\{k - \beta p_L \times \left(1 - F\left(q_{1p}^*\right)\right)\right\} \end{aligned}$$

The difference is positive because $\beta p_L \left(1 - F\left(q_{1p}^*\right)\right) < k$. This leads us to conclude that there is welfare gain for the type L workers by this redistribution.

[4] If the type is public information, this transfer could be made type-dependent, and does not have first order gain.

(2). The increase of first-period unemployment benefit raises type H's welfare by

$$\begin{aligned}\triangle_{H,gain}^{AI} &= \left(1-F\left(q_H^{AI}\right)\right)\times u^{'}(z)\times kdz \\ &\quad +\delta\left[\left(1-F\left(q_L^{*}\right)\right)-\left(1-F\left(q_H^{AI}\right)\right)\right]\times u^{'}(z)\times kdz\end{aligned}$$

where δ is the Lagrangian multiplier of the incentive constraint.[5] Notice that the second term is always greater than zero, because $F(q)$ is a decreasing function and $q_L^{*}>q_H^{AI}$. At the same time, the decrease of second-period z brings type H welfare loss by an amount

$$\begin{aligned}\triangle_{H,loss}^{AI} &= \beta p_H\left(1-F\left(q_H^{AI}\right)\right)\left(1-F\left(q_{1p}^{*}\right)\right)\times u^{'}(z)\,dz \\ &\quad +\delta\beta p_L\left[\left(1-F\left(q_L^{*}\right)\right)\left(1-F\left(q_{1p}^{*}\right)\right)-\left(1-F\left(q_H^{AI}\right)\right)\left(1-F\left(q_{1p}^{*}\right)\right)\right]\times u^{'}(z)\,dz\end{aligned}$$

The second term in this expression is also greater than zero for the same reason just mentioned.

The first term in $\triangle_{H,gain}^{AI}$ is smaller than the first term in $\triangle_{H,loss}^{AI}$, because of the inequality $k<\beta p_H\left(1-F\left(q_{1p}^{*}\right)\right)$. The second term in $\triangle_{H,gain}^{AI}$ is greater than the second term in $\triangle_{H,loss}^{AI}$, because of the inequality $\beta p_L\left(1-F\left(q_{1p}^{*}\right)\right)<k$. At this stage, it is still ambiguous whether it is the gain or loss that dominates. We then make the difference $\triangle_{H,gain}^{AI}-\triangle_{H,loss}^{AI}$, and notice that according to the definition of k we are able to reduce it to the following comparison:

$$\begin{aligned}&\text{Sign}\left[\triangle_{H,gain}^{AI}-\triangle_{H,loss}^{AI}\right] \\ =&\ \text{Sign}\left[\begin{array}{c}\left(1-F\left(q_H^{AI}\right)\right)\times\left(k-\beta p_H\left(1-F\left(q_{1p}^{*}\right)\right)\right) \\ +\delta\left[F\left(q_H^{AI}\right)-F\left(q_L^{*}\right)\right]\times\left(k-\beta p_L\left(1-F\left(q_{1p}^{*}\right)\right)\right)\end{array}\right] \\ =&\ \text{Sign}\left[\begin{array}{c}-\left(1-F\left(q_H^{AI}\right)\right)\times(1-\alpha)\left(1-F\left(q_L^{*}\right)\right) \\ +\delta\left[F\left(q_H^{AI}\right)-F\left(q_L^{*}\right)\right]\times(\alpha)\left(1-F\left(q_H^{AI}\right)\right)\end{array}\right]\end{aligned}$$

Then when $\alpha\to1$, that is, the fraction of type H is high, there will be welfare gain for the type H workers; when $\alpha\to0$, there is welfare loss for the type H workers.

(3). Concerning the variation of the overall welfare, we have

$$\begin{aligned}&\alpha\times\left[\triangle_{H,gain}^{AI}-\triangle_{H,loss}^{AI}\right]+(1-\alpha)\times\left[\triangle_{L,gain}^{AI}-\triangle_{L,loss}^{AI}\right] \\ =&\ \alpha\delta\left[\left(1-F\left(q_L^{*}\right)\right)-\left(1-F\left(q_H^{AI}\right)\right)\right]\times\left[k-\beta p_L\left(1-F\left(q_{1p}^{*}\right)\right)\right]\times u^{'}(z)\,dz\end{aligned}$$

So that

[5] Bi and Langot (2015) show that $0<\delta<1$.

$$\begin{aligned}
& \text{Sign}\left\{\alpha \times \left[\triangle^{AI}_{H,gain} - \triangle^{AI}_{H,loss}\right] + (1-\alpha) \times \left[\triangle^{AI}_{L,gain} - \triangle^{AI}_{L,loss}\right]\right\} \\
= \quad & \text{Sign}\left\{\underbrace{\alpha\delta\left[(1-F(q^*_L)) - \left(1-F\left(q^{AI}_H\right)\right)\right]}_{>0} \times \underbrace{\left[k - \beta p_L \left(1-F\left(q^*_{1p}\right)\right)\right]}_{>0}\right\} \\
= \quad & +
\end{aligned}$$

Then the aggregate welfare increases. At last, we remark that both the direction of type L's welfare variation and the aggregate welfare variation are dependent upon the sign of $\left[k - \beta p_L \left(1-F\left(q^*_{1p}\right)\right)\right]$. Q.E.D.

This proposition shows that if the "distance to retirement" is measured by an exogenous probability of remaining participative (private information), then the type L workers are always the net beneficiary and such a transfer always improves the aggregate welfare; however, the type H workers may lose or gain depending on effects along two dimensions: (1) informational externality, and (2) intensity of cross-subsidization.

By informational externality, we mean that such a transfer of unemployment benefit can relax the informational constraint. To see this, we focus on the incentive constraint: $F(q_H)[u(w_{H,1}) + \beta p_L u(w_{H,2})] + (1-F(q_H))\left[u(z_1) + \beta p_L U_{1p}\right] \leq U_L$. A transfer of unemployment benefit from the old to the young will increase the left-hand side by an amount $\left(1-F\left(q^{AI}_H\right)\right)\left(k - \beta p_L\left(1-F\left(q^*_{1p}\right)\right)\right) u'(z)\,dz$. At the same time, such a redistribution also increases by an amount $(1-F(q^*_L))\left(k - \beta p_L\left(1-F\left(q^*_{1p}\right)\right)\right) u'(z)\,dz$ at the right-hand side. Since Bi and Langot (2015) show that "$q^{AI}_H < q^{FI}_L$ and $F(q)$ is a decreasing function", the right-hand side increases faster than the left-hand side, which implies that the information constraint is easier to be satisfied. This suggests that type L workers become comparatively less willing to misrepresent their type. And we notice in the end of proof Proposition 1 that the aggregate welfare is enhanced, and the magnitude of that enhancement is solely related to the information gain which has δ as a multiplier.

We now turn to cross-subsidization. There is heterogeneity in the population: type L and type H. On one hand, the fact that unemployed type H workers have higher probability (p_H) of remaining in the market tomorrow leads them to be sampled more frequently by the planner for the collection of unemployment benefit as a source of redistribution. On the other hand, the fact that they are matched more often in the first period suggests that they are less likely to be unemployed in the first period and hence less likely to benefit from any transfer. A decrease of value dz then will be more likely to be redistributed to the type L workers - type H will cross-subsidize type L. Then pure cross-subsidization represents a welfare loss to type H workers, albeit relaxed information constraint tends to raise their welfare. When the fraction α of type H workers is sufficiently low, it implies that a small fraction of type H will cross-subsidize a large fraction of type L. Then the gain from reduction of information externality (captured by the term with δ) is insufficient compared to the excessive loss from cross-subsidization[6], type H workers are subsequently worse off. On the other hand, if the fraction α of type H workers is sufficiently high, a large fraction of type H will take care of a small fraction of type L. Then the loss from cross-subsidization is dominated by the information gain, so that there is welfare gain for type H.

2.3 Endogenous participation choice before search and age-specific UB

In the previous section, workers' premature quit is governed by an exogenous staying probability p_{type}. However, to stay in the market for longer or shorter periods is often an idiosyncratic decision. Bi and Langot (2015) make a simple modification to the previous model to enable the possibility that workers make ex ante decisions on submarkets (distinguished by employment horizon) choice, so that the adjustments through the extensive margin become possible. Bi and Langot (2015) show that distortion induced by self-selection process can dampen the participation incentive of workers with lower value of leisure. In this section, we wish to examine the effectiveness of age-specific unemployment benefits under such a context.

2.3.1 Equilibrium

The submarket choice is endogenous in the following sense: each agent receives an idiosyncratic signal (private information) at the ex ante stage on her home production value (denoted h) for the second period. Given this individual characteristic, she chooses ex ante to join the two-period or the one-period submarket. According to Bi and Langot (2015), the two submarkets are separated by one threshold $\hat{h}$. For values of h below $\hat{h}$, the worker chooses to join the two-period submarket: she will then quit with probability 0, and this worker will be categorized as type H. Otherwise, the employment relationship is only of one period,

[6]When α decreases 1 percent, $(1-\alpha)$ increases 1 percent. But their difference in absolute value increases 2 percents.

and the worker will be categorized as type L. Formally, if h is a continuous variable from a distribution $G(h)$ over the support $\left[\underline{h}, \overline{h}\right]$, we establish that *(i)* among the pool of type L workers, there is no information problem; *(ii)* among the pool of type H workers, there is also no information problem; *(iii)* Type L workers will misrepresent their types as type H workers. The threshold $\hat{h}$ is determined by comparing the expected utility from joining each submarket. With full information, this threshold is determined by $U_L^{FI}\left(\hat{h}^{FI}\right) = U_H^{FI}\left(\hat{h}^{FI}\right)$, and with asymmetric information, it is determined by $U_L\left(\hat{h}^{AI}\right) = U_H^{AI}\left(\hat{h}^{AI}\right)$. We derive both situations and make comparisons.

The maximization problem for the firms in the two-period submarket is:

$$\begin{array}{rcl} U_H^{AI}(h_H, z_1, z_2) & = & \max_{q, w_{H,1}, w_{H,2}} \left\{ \begin{array}{c} F(q_H)\left[u(w_{H,1}) + \beta u(w_{H,2})\right] \\ +(1-F(q_H))\left[u(z_1) + \beta \max\left\{u(h_H), U_{1p}(h_H, z_2)\right\}\right] \end{array} \right\} \\ \text{s.t.} & & H(q_H)\left[y - w_{H,1} + \beta(y - w_{H,2})\right] = C \\ IC \text{ constraint} & & F(q_H)\left[u(w_{H,1})\right] + (1-F(q_H))\left[u(z_1)\right] + \beta u(h_L) \leq U_L^* \end{array}$$

where the solution for the contract posted by firms that target one-period workers is given by

$$\begin{array}{rcl} U_L^*(h_L, z_1) & = & \max_{q, w_L}\left\{F(q)u(w_L) + (1-F(q))u(z_1) + \beta u(h_L)\right\} \\ \text{s.t.} & & H(q)\left[y - w_L\right] = C \end{array}$$

Finally, the optimal contract posted by firms which target workers that prefer to search for a job at the second period is a solution to the following problem:

$$\begin{array}{rcl} U_{1p}(h_H, z_2) & = & \max_{q, w}\left\{F(q)u(w) + (1-F(q))\left[u(h_H) + u(z_2)\right]\right\} \\ \text{s.t.} & & H(q)\left[y - w\right] = C \end{array}$$

The optimal contract is such that *(i)* the wages satisfy $w_{H,1}^{AI} < w_H^{FI} < w_{H,2}^{AI}$, *(ii)* the queue lengths are such that $q_H^{FI} < q_H^{AI} < q_L^* = q_L^{FI}$, and *(iii)* the submarket choice leads to a threshold $\hat{h}^{AI} < \hat{h}^{FI}$. With asymmetric information, firms in the two-period market must discourage the type L workers from misrepresenting their types, and again type L workers receive information rent. Distorted allocation for type H decreases their expected utility, leading some of them to choose non-participation. Hence when ex ante decision on participation is allowed, the distorted contracts lower the aggregate labor supply in the presence of asymmetric information.

2.3.2 Age-Specific UB

As previously, the policy consists of front-loading redistributing unemployment benefit by age through increasing the first-period z by an amount dz, and decreasing the second-period z by dz. The government will expect to gather money from the following fraction of the old:

$$G\left(\hat{h}^i\right) \times \left[\beta\left(1-F\left(q_H^i\right)\right)\left(1-F\left(q_{1p}^*\right)\right)\right] \tag{2.3}$$

where $G\left(h^i\right)$ is the fraction of type H workers who remain in the labor market tomorrow, and $i \in \{AI, FI\}$. At the same time, the government will expect to redistribute to the following fraction of current generation who suffers from unemployment:

$$G\left(\hat{h}^i\right) \times \left(1-F\left(q_H^i\right)\right) + \left(1-G\left(h^i\right)\right) \times \left(1-F\left(q_L^i\right)\right) \tag{2.4}$$

The expected fraction of workers who receive the first-period $u(z_1)$ are those who are not able to find a job in the first period, and this fraction consists of $\left(1-F\left(q_H^i\right)\right)$ for type H, and $\left(1-F\left(q_L^i\right)\right)$ for type L. The expected fraction of workers who receive the second-period $u(z_2)$ consists of $\beta\left(1-F\left(q_H^i\right)\right)\left(1-F\left(q_{1p}^*\right)\right)$ for type H, and there are no type L workers who are in the market tomorrow. Now to keep the budget balanced, every unit of decrease of old generation's unemployment benefit dz should be translated to

$$\tilde{k} \equiv \frac{G\left(\hat{h}^i\right) \times \left[\beta\left(1-F\left(q_H^i\right)\right)\left(1-F\left(q_{1p}^*\right)\right)\right]}{G\left(\hat{h}^i\right) \times \left(1-F\left(q_H^i\right)\right) + \left(1-G\left(\hat{h}^i\right)\right) \times \left(1-F\left(q_L^i\right)\right)} \times dz$$

unit of utility gain for the young generation regardless of the type. Such a type-independent transfer is justified by the fact that the planner has also no information on the agents' types. It could be straightforwardly verified that we have $\frac{\tilde{k}}{\left(1-F\left(q_{1p}^*\right)\right)} < \beta$.

Proposition 2.

(1) Front-loading redistribution of unemployment benefit by age has the same impact on welfare as summarized in Proposition 1.

(2) After redistribution, the participation margin for both full information ($\hat{h}^{FI}$) and asymmetric information ($\hat{h}^{AI}$) decrease.

Proof. We start by (1). By Envelope Theorem, the increase of first-period z brings type L workers welfare gain by an amount of $\triangle_{L,gain}^i = \left(1-F\left(q_L^i\right)\right) \times u'(z) \times \tilde{k}dz$. At the same time, the decrease of second-period z does not concern type L workers. Then there is net welfare gain for the type L workers by this unemployment insurance redistribution. Now we turn to type H workers.

This increase of first-period unemployment benefit raises type H workers' welfare by

$$\begin{aligned}\triangle^i_{H,gain} &= \left(1-F\left(q^i_H\right)\right)\times u^{'}(z)\times\tilde{k}dz \\ &+\delta\left[\left(1-F\left(q^i_L\right)\right)-\left(1-F\left(q^i_H\right)\right)\right]\times u^{'}(z)\times\tilde{k}dz\end{aligned}$$

where $\delta<1$ is the Lagrangian multiplier of the incentive constraint. Notice that the second term is always greater than zero. At the same time, the decrease of second-period z brings type H workers welfare loss by an amount

$$\triangle^i_{H,loss} = \beta\left(1-F\left(q^i_H\right)\right)\left(1-F\left(q^*_{1p}\right)\right)\times u^{'}(z)\,dz$$

The first term in $\triangle^i_{H,gain}$ is smaller than the first term in $\triangle^i_{H,loss}$, because of the inequality $\frac{\tilde{k}}{\left(1-F\left(q^*_{1p}\right)\right)}<\beta$. The second term in $\triangle^i_{H,gain}$ is strictly greater than zero. At this stage, it is still ambiguous whether it is the gain or loss that dominates. We then make the difference $\triangle^i_{H,gain}-\triangle^i_{H,loss}$, and notice that according to the definition of k we are able to reduce it to the following comparison:

$$\begin{aligned}&\text{Sign}\left[\triangle^i_{H,gain}-\triangle^i_{H,loss}\right] \\ =\ &\text{Sign}\left[\begin{array}{c}-\left(1-G\left(\hat{h}^i\right)\right)\left(1-F\left(q^i_L\right)\right) \\ +\delta\left[F\left(q^i_H\right)-F\left(q^i_L\right)\right]\times G\left(\hat{h}^i\right)\end{array}\right]\end{aligned}$$

Again, when $G\left(h^i\right)$ is such that the fraction of type H is high, there will be welfare gain for the type H workers; when $G\left(h^i\right)$ is such that the fraction of type H is low, there will be welfare loss for the type H workers.

Concerning the variation of the aggregate welfare, we have

$$\begin{aligned}&G\left(\hat{h}^i\right)\times\left[\triangle^i_{H,gain}-\triangle^i_{H,loss}\right]+\left(1-G\left(\hat{h}^i\right)\right)\times\left[\triangle^i_{L,gain}-\triangle^i_{L,loss}\right] \\ =\ &\delta\left[\left(1-F\left(q^i_L\right)\right)-\left(1-F\left(q^i_H\right)\right)\right]\times\tilde{k}G\left(\hat{h}^i\right)\times u^{'}(z)\,dz \\ >\ &0\end{aligned}$$

Then the aggregate welfare increases after redistribution. Thus we obtain similar result as before, except that now the fraction is no longer α, but is $G\left(\hat{h}^i\right)$ which depends on the distribution function $G(.)$ and $\hat{h}^i$, the value of which is determined in the equilibrium.

We now investigate (2): how this front-loading redistribution of unemployment benefit by age has impact on $\hat{h}$ which is determined by the indifference condition $U_L\left(\hat{h}^{FI}\right)=U^{FI}_H\left(\hat{h}^{FI}\right)$ for the case of full information and $U_L\left(\hat{h}^{AI}\right)=U^{AI}_H\left(\hat{h}^{AI}\right)$ for the case of asymmetric information. We start with the case of full information. We established above that U_L^{FI} increases but U_H^{FI} decreases due to intergenerational redistribution of z. Then given that both U_L^{FI} and U_H^{FI} are increasing function of h by Envelope Theorem, $\hat{h}^{FI}$ must fall after the intergenerational transfer takes place to make the equality stay binding. In case of asymmetric information, although depending on the value of $G\left(\hat{h}^i\right)$ the value of U_H^{AI} may increase, it can be straightforwardly checked that $\triangle_{H,gain}^{AI} - \triangle_{H,loss}^{AI} < \triangle_{L,gain}^{AI} - \triangle_{L,loss}^{AI}$, which says that the increase of U_L^{FI} is more important compared to increase of U_H^{AI}, so that again $\hat{h}^{AI}$ must fall to make the equality stay binding after the intergenerational transfer takes place. Q.E.D.

This proposition suggests that there is always aggregate welfare gain from this redistribution, however, the threshold $\hat{h}$ is downwards distorted: there are more workers who become reluctant to choose long-horizon employment relationship in the two-period submarket, and are willing to switch to the one-period market. We suggest that this situation implies potential loss in the (economy-wide) aggregate output (the sum of aggregate output in the second-period and one-period submarket). To proceed heuristically our argument, we follow literature to propose the measure for the expected aggregate output in the respective submarket as "expected number of formed matches"[7] times "the ex post output" (y for the one-period and $(1+\beta)y$ for the two-period market). Given this measure, when workers switch from two-period to one-period submarket, there is a discrete drop on the ex post surplus from $(1+\beta)y$ to y, the magnitude of which depends on β; At the same time if the decrease of "expected number of matches" in second-period submarket is not too small compared to the increase of "expected number of matches" in the one-period submarket, then in total the reduction of output in the former is more important compared to corresponding gain in the latter - there is hence economy-wide aggregate output loss.

[7]The "expected number of matches" is also called a matching function $m(N,M)$, which depends on the number of unemployed seekers for job (of number N) and number of vacant firms (of number M), is increasing in both argument, and homogenous in degree 1. In our context, with $q = \frac{N}{M}$, we have $H(q) = \frac{m(N,M)}{M} = m\left(\frac{N}{M}, 1\right)$, and $F(q) = \frac{m(N,M)}{N}$. Recall that N is exogenous and M endogenously determined by free entry - zero profit. We refer readers to p392 of the survey provided by Petrongolo & Pissarides (2001) for more details on matching functions.

2.4 Ex-post asymmetric information, participation choices after matching and age-specific UB

2.4.1 The model

In this section, we consider the situation where ex ante identical workers become informed about their private value of home production only after getting matched, and such a shock takes place at the end of the first period during this employment relationship. Depending on the realization of this home production level, the worker decides to stay for another period or to quit the labor market early (without paying any compensation to the firm). We assume that workers' private information on the home production shock h is drawn randomly from a cumulative distribution function $G(h)$, with the support $\left[\underline{h},\overline{h}\right]$. When the value of h is high, workers are better off quitting the job to enjoy it. Thus the firm's problem is to propose a wage contract that *(i)* guarantees participation (Individual Rationality or I.R.), and *(ii)* makes sure that workers reveal truthfully their realisation of h (Incentive Compatibility or I.C.). By standard mechanism design approach, Bi and Langot (2015) reduce the firm's maximization program to the following one:

$$
\begin{array}{llr}
U^{AI} = & \max\limits_{w_1(h),w_2(h),e(h),q} & F(q)\times\left[u(w_1)+\beta\int_{\underline{h}}^{\overline{h}}\left\{\begin{array}{c} e(h)\,u(w_2(h)) \\ +(1-e(h))\,u(h)\end{array}\right\}dG(h)\right] \\
 & & +(1-F(q))\left\{u(z)+\beta\int_{\underline{h}}^{\overline{h}}\max\left\{U_{1p}(h,z),u(h)\right\}dG(h)\right\} \\
s.to & \text{resource constraint} & F(q)\int_{\underline{h}}^{\overline{h}}\left[y-w_1+\beta e(h)\times(y-w_2)\right]dG(h)-\frac{1}{q}K=0 \\
\text{and} & \text{I.C. + I.R.} & \int_{\underline{h}}^{\overline{h}}e(h)\left[u(w_2)-u(h)-\frac{G(h)}{g(h)}u'(h)\right]dG(h)\geq 0
\end{array}
$$

In this program, $e(h)\in\{0,1\}$ will stand for the probability with which the firm decides to retain the workers, and we have

$$
\begin{array}{rcl}
U_{1p}(h,z) & = & \max\limits_{w} F\left(q_{1p}\right)u\left(w_{1p}\right)+\left(1-F\left(q_{1p}\right)\right)\left[u(h)+u(z)\right] \\
\text{s.to} & & H\left(q_{1p}\right)\left(y-w_{1p}\right)\geq C
\end{array}
$$

Also notice that we assume the utility between the home production h and unemployment benefit to be separable: that is $u(h+z)=u(h)+u(z)$. Bi and Langot (2015) characterized the equilibrium properties as follows: *(i)* the wage profile is such that $w_1^{AI}<w_2^{AI}<w^{FI}$, *(ii)* the optimal queue length is such that $q^{AI}>q^{FI}$, *(iii)* the quitting threshold is such that $\hat{h}^{AI}<\hat{h}^{FI}$. Although the first best (full information) allocation can not be achieved, the constrained efficiency can be obtained. That social planner will maximize the workers' expected utility from search subject to the following constraints: *(1)* the resource constraint which stipulates that the expected aggregate consumption should not surpass the expected aggregate output; *(2)* the individual rationality constraint which requires that the workers should have incentive to participate and search for the job; *(3)* given that z is fixed and can not be redistributed (Guerrieri (2008) terms it nontransferrable), the consumption for unemployment workers should be at least as great as z: that is to say, $Z_u \leq z$.

2.4.2 Redistribution by age.

It is straightforward to observe that since the unemployment benefit z does not enter into the information constraint, there is no welfare gain in the first order from a smaller amount of redistribution. We now show that in case the workers who quit prematurely have the right to claim for the unemployment benefit, there is still welfare improvement from front-loading unemployment benefit. Such a redistribution also raises the participation margin.

Proposition 3. *In the case of ex post heterogeneity, redistributing unemployment benefit by age relaxes information constraint.*

Proof. The result that such a transfer brings the improvement of welfare is actually quite natural, because in the current context, there is no effect from cross-subsidization, and such a transfer simply relaxes the information constraint, and thus brings improvement to the welfare. To see this, we consider the following program where we distinguish the first-period unemployment benefit by denoting it z_1 and the second-period unemployment benefit by labeling it as z_2.

$$\begin{array}{lcl}
U^{AI}(z_1,z_2) = & \max\limits_{w_1(h),w_2(h),e(h),q} & F(q)\int_{\underline{h}}^{\overline{h}} \left[u(w_1) + \beta \left[\begin{array}{cc} & e(h)\,u(w_2(h)) \\ + & (1-e(h))\left[u(h)+u(z_2)\right] \end{array} \right] \right] dG(h) \\
& & +(1-F(q))\left\{ u(z_1) + \beta \int_{\underline{h}}^{\overline{h}} \max\left\{U_{1p}(h,z_2), u(h)\right\} dG(h) \right\} \\
s.to & \text{Resource} & F(q)\int_{\underline{h}}^{\overline{h}} \left[y - w_1 + \beta e(h)\times(y-w_2)\right] dG(h) - \frac{1}{q}K = 0 \\
\text{and} & \text{Information} & \int_{\underline{h}}^{\overline{h}} e(h)\left[u(w_2) - u(h) - u(z_2) - \frac{G(h)}{g(h)}u'(h)\right] dG(h) \geq 0
\end{array}$$

By Envelope Theorem, an increase of dz for the young workers will bring a welfare gain of the level

$$\overbrace{\left(1-F\left(q^{AI}\right)\right)}^{\text{A1}} u'(z) \times \check{k}dz$$

At the same time, a decrease of dz for the old workers (at the second period of the lifetime) will bring a welfare loss of the magnitude

$$
\overbrace{\beta F\left(q^{AI}\right)\left(1-G\left(\hat{h}^{AI}\right)\right)}^{\text{A2}} \times u^{'}(z)\,dz
$$

$$
+\left[\underbrace{\beta\left(1-F\left(q^{AI}\right)\right)\int_{\underline{h}}^{\overline{h}}\left[e_I(h)\left(1-F\left(q_I\right)\right)+\left(1-e_I(h)\right)\right]dG(h)}_{\text{A3}}-\underbrace{\delta\times G\left(\hat{h}^{AI}\right)}_{\text{A4}}\right]\times u^{'}(z)\,dz
$$

where $e_I(h)$ determines the threshold captured by max $\left\{U_{1p}(h,z),u(h)\right\}$. Analogous as before, $\breve{k}$ is defined such that the we have budget balance, that is to say, $\text{A1}\times\breve{k}=\text{A2}+\text{A3}-\text{A4}$.Given such a relationship, the aggregation of the above gain and loss becomes $\text{A1}\times\breve{k}-(\text{A2 + A3}-\text{A4})=\text{A4}>0=\delta\times G\left(\hat{h}^{AI}_{II}\right)$. Thus, we have shown that such a transfer has a first order effect on the improvement of welfare, and such an improvement comes solely from the reduction of information externality.

Bi and Langot (2015) suggest that δ depends negatively on $\hat{h}^{AI}$. When information constraint is relaxed, δ becomes smaller, so that the value of $\hat{h}^{AI}$ naturally increases towards $\hat{h}^{FI}$. Q.E.D.

We should make some comparisons between the current context and the previous contexts. The first difference between the current context with ex post heterogeneity and the previous contexts with ex ante heterogeneity is the absence of cross-subsidization. If there were cross-subsidization, some types necessarily benefit more than average and some benefit less than average, and the low types always gain due to cross-subsidization from the high types. Improving the low type's welfare makes them less willing to misrepresent their type, so that there is welfare improvement. The second difference comes from the comparison between the current context and that of ex ante heterogeneity with submarket choice. With ex ante heterogeneity, such a transfer by age always further distorts the participation margin downwards due to the faster increase of welfare of the low type, this has potentially negative effect on the aggregate output. However, with ex post heterogeneity, such a transfer has an effect of increasing the participation margin which positively affects the aggregate output. This is because reducing second-period unemployment benefit makes quitting a less attractive option. Apart from the above two main differences, all the above contexts predict that such redistributive transfer by age reduces the information rent.

2.5 Discussion on the robustness of the results

In the previous two-period contexts, the main reason for overall welfare improvement is that whenever we front-load UB from the second to the first period, the incentive constraint is relaxed. At the same time, type L workers benefit from cross-subsidization so that their welfare is always raised. In this section, we move to a three-period extension of the benchmark proposed in the discussion section of Bi and Langot (2015).[8] We examine whether the results achieved in the previous section can carry through. In the following, the notations "l", "m", and "s" will represent respectively Long-term, Middle-term, Short-term employment relationship. The superscript "AI" will be given to the programs with incentive constraints. The solution of the program without information constraints will be labeled by the superscript "$*$". Bi and Langot (2015) suggest that the results on equilibrium allocations of the benchmark could be carried through. The program is

$$\begin{array}{rcl} U_{l,1}^{AI}(H) & = & \max_{w_{l,1},w_{l,2},w_{l,3},q_{l,1}} \left\{ \begin{array}{c} F\left(q_{l,1}\right)\left[u\left(w_{l,1}\right)+\beta p_H\left[u\left(w_{l,2}\right)+\beta p_H u\left(w_{l,3}\right)\right]\right] \\ +\left(1-F\left(q_{l,1}\right)\right)\left[u\left(z_1\right)+\beta p_H U_{l,2}^{AI}(H)\right]+\beta\left(1-p_H\right)\times\left[\beta p_H U_{l,3}^{*}\right] \end{array} \right\} \\ \text{s.to} & & H\left(q_{l,1}\right)\left[y-w_{l,1}+\beta p_H\left(y-w_{l,2}\right)+\beta^2 p_H^2\left(y-w_{l,3}\right)\right]=C \\ & & F\left(q_{l,1}\right)\left[u\left(w_{l,1}\right)+\beta p_L\left[u\left(w_{l,2}\right)+\beta p_L u\left(w_{l,3}\right)\right]\right] \\ & & +\left(1-F\left(q_{l,1}\right)\right)\left[u\left(z_1\right)+\beta p_L U_{l,2}^{*}(L)\right]+\beta\left(1-p_L\right)\times\left[\beta p_L U_{l,3}^{*}\right]\leq U_{l,1}^{*}(L) \end{array}$$

It is important to notice that there is a particular term in the program: $\beta\left(1-p_t\right)\times\left[\beta p_t U_{l,3}^{*}\right]$ where $t\in\{L,H\}$. This term comes from the fact that the workers who are inactive at the second period (captured by $\beta\left(1-p_t\right)$) could be again active in the market at the third period (captured by βp_t). The presence of this term renders our analysis of individual type's welfare ambiguous, hence we will put focus on the aggregate welfare, and will suggest in the Appendix to which extent this term modifies the result. To proceed, there are three situations to distinguish according to the source and the destination of the redistribution. In the first situation (Situation 1), we front-load unemployment benefit from the third to the second period; In the second situation (Situation 2), the transfer is from the third to the first period; In the third situation (Situation 3), it is from the second to the first period.

Proposition 4.

(1) In Situation 1, the aggregate welfare always improves after front-loading unemployment benefit.

(2) In Situation 2 and Situation 3, a sufficient condition for the aggregate welfare to improve is $p\times\left(1-F\left(q_{m,1}^{*}(p)\right)\right)$ *being increasing in* p.

Proof. In the Appendix.

[8] Due to the complexity of the analysis, we mainly focus on the extension of the benchmark here.

In Appendix, it is shown that the $(1-p)\times p$ effect makes the above result on aggregate welfare gain stronger, because this effect brings additional source for the funding. Absent this effect, we find that the welfare variation depends crucially on the comparison between $p_H\left(1-F\left(q_{m,1}^{AI}(H)\right)\right)$ and $p_L\left(1-F\left(q_{m,1}^{*}(L)\right)\right)$ as shown in the proof, and this comparison is involved in Situation 2 and Situation 3. In Situation 3, it is related because the asymmetric information on p for the third period distorts the matching probability for the second period; and the relevant comparison is on $\beta p_H\left(1-F\left(q_{m,1}^{AI}(H)\right)\right)$ and $\beta p_L\left(1-F\left(q_{m,1}^{*}(L)\right)\right)$. In Situation 2, this comparison becomes $\beta p_H\left(1-F\left(q_{m,1}^{AI}(H)\right)\right)\times\beta p_H\left(1-F\left(q_{1p}^{*}\right)\right)$ and $\beta p_L\left(1-F\left(q_{m,1}^{*}(L)\right)\right)\times\beta p_L\left(1-F\left(q_{1p}^{*}\right)\right)$, simply because the distortion for the second third period is accumulated to the third period. And it is intuitive that in Situation 1, the comparison rests on $\beta p_H\left(1-F\left(q_{1p}^{*}\right)\right)$ and $\beta p_L\left(1-F\left(q_{1p}^{*}\right)\right)$, which renders no ambiguity. On one hand, $p_H > p_L$: type H workers have higher probability of remaining in the market; on the other hand, $\left(1-F\left(q_{m,1}^{AI}(H)\right)\right) < \left(1-F\left(q_{m,1}^{*}(L)\right)\right)$: it is less probable that type H become unemployed; the comparison between $p_H\left(1-F\left(q_{m,1}^{AI}(H)\right)\right)$ and $p_L\left(1-F\left(q_{m,1}^{*}(L)\right)\right)$ is hence ambiguous. At the same time, according to Envelope Theorem, these expressions also correspond respectively to $\frac{dU_{m,1}^{AI}(H)}{dU_{1p}}$ and $\frac{dU_{m,1}(L)}{dU_{1p}}$, which represent the marginal gain from one unit of increase of the value U_{1p} related to another possibility of search in the subsequent period respectively for type H and type L. Since $\left(1-F\left(q_{m,1}^{AI}(H)\right)\right) > \left(1-F\left(q_{m,1}^{*}(H)\right)\right)$, our assumption that "$p\times\left(1-F\left(q_{m,1}^{*}(p)\right)\right)$ being increasing in p" hence requires that in case of full information a unit increase of U_{1p} is always more valuable for the type H workers. Then conversely, a unit decrease of U_{1p} costs for type H workers more compared to type L, hence cross-subsidization has a direction from type H towards type L. Based on these descriptions, we think this is a strong assumption.

2.6 Conclusion

In this paper, we analyze the welfare impact of front-loading unemployment benefit by age redistributively. In the two-period benchmark, we find that there is always aggregate welfare gain from redistribution and the type L workers always benefit from cross-subsidization. We then allow the workers to choose their horizon. Indeed, when the information is realized before search, this transfer triggers comparatively larger welfare improvement for the potential one-period workers, which hence lowers the participation margin, and in turn may bring down aggregate output; in contrast, when the participation shock is realized after matching, this redistribution always increases participation margin and the aggregate output. We also check the robustness of the benchmark in the three-period setup, and argue that the aggregate welfare gain could be achieved under a fairly strong condition.

2.7 Appendix

Proposition 4.

(1) In Situation 1, the aggregate welfare always improves after front-loading unemployment benefit.

(2) In Situation 2 and Situation 3, a sufficient condition for the aggregate welfare to improve is $p\times\left(1-F\left(q^*_{m,1}(p)\right)\right)$ being increasing in p.

Proof. We start by Situation 1: Front-loading unemployment benefit from the third to the second period. The fraction of third-period type H workers who receive unemployment benefit is

$$\alpha\times\left\{\begin{array}{lc} & \left(1-F\left(q^{AI}_{l,1}(H)\right)\right)\times\left[\beta p_H\times\left(1-F\left(q^{AI}_{m,1}(H)\right)\right)\times\beta p_H\left(1-F\left(q^*_{s,1}\right)\right)\right] \\ + & \left[\beta p_H\times\beta\left(1-p_H\right)\left(1-F\left(q^*_{s,1}\right)\right)\right]\end{array}\right\}$$

The fraction of third-period type L workers who receive unemployment benefit is

$$\begin{array}{c}(1-\alpha)\times \\ \left\{\begin{array}{lc} & \left(1-F\left(q^{*}_{l,1}(L)\right)\right)\times\left[\beta p_L\times\left(1-F\left(q^{*}_{m,1}(L)\right)\right)\times\beta p_L\left(1-F\left(q^*_{s,1}\right)\right)\right] \\ + & \left[\beta p_L\times\beta\left(1-p_L\right)\left(1-F\left(q^*_{s,1}\right)\right)\right]\end{array}\right\}\end{array}$$

The fraction of second-period type H workers who receive unemployment benefit is

$$\alpha\left(1-F\left(q^{AI}_{l,1}(H)\right)\right)\times\beta p_H\times\left(1-F\left(q^{AI}_{m,1}(H)\right)\right)$$

The fraction of second-period type L workers who receive unemployment benefit is

$$(1-\alpha)\left(1-F\left(q^{*}_{l,1}(L)\right)\right)\times\beta p_L\times\left(1-F\left(q^{*}_{m,1}(L)\right)\right)$$

To keep budget balanced, every unit of third period unemployment benefit is translated to $\bar{k}_{32}$ unit for the second period unemployed. So that we have

$$\bar{k}_{32} = \frac{\left\{\begin{array}{c}\alpha\times\beta p_H\left(1-F\left(q^*_{s,1}\right)\right)\left\{\beta p_H\times\left(1-F\left(q^{AI}_{l,1}(H)\right)\right)\left(1-F\left(q^{AI}_{m,1}(H)\right)\right)+\beta\left(1-p_H\right)\right\} \\ (1-\alpha)\times\beta p_L\left(1-F\left(q^*_{s,1}\right)\right)\left\{\beta p_L\times\left(1-F\left(q^{*}_{l,1}(L)\right)\right)\left(1-F\left(q^{*}_{m,1}(L)\right)\right)+\beta\left(1-p_L\right)\right\}\end{array}\right.}{\left[\begin{array}{c}\alpha\left(1-F\left(q^{AI}_{l,1}(H)\right)\right)\times\beta p_H\times\left(1-F\left(q^{AI}_{m,1}(H)\right)\right) \\ +(1-\alpha)\left(1-F\left(q^{*}_{l,1}(L)\right)\right)\times\beta p_L\times\left(1-F\left(q^{*}_{m,1}(L)\right)\right)\end{array}\right]}$$

The next step is to determine how does such a redistribution of a small magnitude dz alternate the welfare of type L workers, that is, we should find the value of $\triangle^{AI}_{L,gain,32}-\triangle^{AI}_{L,loss,32}$. We have

$$\begin{aligned}
& \triangle^{AI}_{L,gain,32} - \triangle^{AI}_{L,loss,32} \\
= & \beta p_L \left(1 - F\left(q^*_{l,1}(L)\right)\right)\left(1 - F\left(q^*_{m,1}(L)\right)\right) \\
& \times \left[\bar{k}_{32} - \left(1 - F\left(q^*_{s,1}\right)\right)\left(\beta p_L + \beta(1-p_L)\frac{1}{\left(1-F\left(q^*_{l,1}(L)\right)\right)\left(1-F\left(q^*_{m,1}(L)\right)\right)}\right)\right] \times u'(z)\,dz
\end{aligned}$$

So that

$$\begin{aligned}
& \text{Sign}\left\{\triangle^{AI}_{L,gain,32} - \triangle^{AI}_{L,loss,32}\right\} \\
= & \ \text{Sign}\left\{\bar{k}_{32} - \left(1 - F\left(q^*_{s,1}\right)\right)\left(\beta p_L + \beta(1-p_L)\frac{1}{\left(1-F\left(q^*_{l,1}(L)\right)\right)\left(1-F\left(q^*_{m,1}(L)\right)\right)}\right)\right\} \\
= & \ \text{Sign}\left\{\begin{array}{ll} & \left(1 - F\left(q^*_{s,1}\right)\right)\left(\beta p_H + \beta(1-p_H)\frac{1}{\left(1-F\left(q^{AI}_{l,1}(H)\right)\right)\left(1-F\left(q^{AI}_{m,1}(H)\right)\right)}\right) \\ - & \left(1 - F\left(q^*_{s,1}\right)\right)\left(\beta p_L + \beta(1-p_L)\frac{1}{\left(1-F\left(q^*_{l,1}(L)\right)\right)\left(1-F\left(q^*_{m,1}(L)\right)\right)}\right) \end{array}\right\}
\end{aligned}$$

This term is ambiguous because of the $(1-p)\,p$ effect. Although $p_H > p_L$, the comparison between $\frac{\beta(1-p_H)}{\left(1-F\left(q^{AI}_{l,1}(H)\right)\right)\left(1-F\left(q^{AI}_{m,1}(H)\right)\right)}$ and $\frac{\beta(1-p_L)}{\left(1-F\left(q^*_{l,1}(L)\right)\right)\left(1-F\left(q^*_{m,1}(L)\right)\right)}$ is ambiguous. On one hand $(1-p_H) < (1-p_L)$, on the other hand, $\frac{1}{\left(1-F\left(q^{AI}_{l,1}(H)\right)\right)\left(1-F\left(q^{AI}_{m,1}(H)\right)\right)} > \frac{1}{\left(1-F\left(q^*_{l,1}(L)\right)\right)\left(1-F\left(q^*_{m,1}(L)\right)\right)}$, because $q^{AI}_{l,1}(H) < q^*_{l,1}(L)$ and $q^{AI}_{m,1}(H) < q^*_{m,1}(L)$ from Bi and Langot (2015). The $(1-p)\,p$ effect is related to the terms with $(1-p_t)$ in the expression. So that the presence of this $(1-p)\,p$ effect could either make cross-subsidization for type L better or worse.

Now we turn to the overall welfare, the improvement of which is measured by

$$dW_{32} = \alpha \times \left(\triangle^{AI}_{H,gain,32} - \triangle^{AI}_{H,loss,32}\right) + (1-\alpha) \times \left(\triangle^{AI}_{L,gain,32} - \triangle^{AI}_{L,loss,32}\right)$$

It is straightforward to show that the resulting expression is only involved with the terms associated with the Lagrangian multiplier of the information constraint:

$$\begin{aligned}
dW_{32} = & \ \alpha \times \left(\triangle^{AI}_{H,gain,32} - \triangle^{AI}_{H,loss,32}\right) + (1-\alpha) \times \left(\triangle^{AI}_{L,gain,32} - \triangle^{AI}_{L,loss,32}\right) \\
= & \ \alpha \times \delta_l \times \left\{\begin{array}{l} + \left(\begin{array}{c} \beta p_L \times \left(1 - F\left(q^*_{l,1}(L)\right)\right)\left(1 - F\left(q^*_{m,1}(L)\right)\right) \\ \times \left[\bar{k}_{32} - \beta p_L \left(1 - F\left(q^*_{s,1}\right)\right) \times 1\right] \end{array}\right) \\ - \left(\begin{array}{c} \beta p_L \times \left(1 - F\left(q^{AI}_{l,1}(H)\right)\right)\left(1 - F\left(q^*_{m,1}(L)\right)\right) \\ \times \left[\bar{k}_{32} - \beta p_L \left(1 - F\left(q^*_{s,1}\right)\right) \times 1\right] \end{array}\right) \end{array}\right\} \times u'(z)\,dz
\end{aligned}$$

Then $\text{Sign}\{dW_{32}\} = \text{Sign}\left\{\left[\bar{k}_{32} - \beta p_L\left(1 - F\left(q^*_{s,1}\right)\right)\right]\right\}$. When the effect of $(1-p)\,p$ is absent, it could be shown that $\text{Sign}\{dW_{32}\}$ is always positive. When the effect of $(1-p)\,p$ is present, the value of $\bar{k}_{32}$ is larger, so that there is always aggregate welfare gain in Situation 1.

We now treat Situation 2: Front-loading unemployment benefit from the third period to the first period. The fraction of first-period type H workers who receive unemployment benefit is

$$\alpha\left(1 - F\left(q^{AI}_{l,1}(H)\right)\right)$$

The fraction of first-period type L workers who receive unemployment benefit is

$$(1-\alpha)\left(1 - F\left(q^*_{l,1}(L)\right)\right)$$

Budget balance requires that every unit of third period unemployment benefit is translated to $\bar{k}_{31}$ unit for the first period unemployed:

$$\bar{k}_{31} = \frac{\left\{\begin{array}{c} \alpha \times \beta p_H\left(1-F\left(q^*_{s,1}\right)\right)\left\{\beta p_H \times \left(1-F\left(q^{AI}_{l,1}(H)\right)\right)\left(1-F\left(q^{AI}_{m,1}(H)\right)\right)+\beta\left(1-p_H\right)\right\} \\ (1-\alpha)\times\beta p_L\left(1-F\left(q^*_{s,1}\right)\right)\left\{\beta p_L \times \left(1-F\left(q^*_{l,1}(L)\right)\right)\left(1-F\left(q^*_{m,1}(L)\right)\right)+\beta\left(1-p_L\right)\right\} \end{array}\right.}{\left[\begin{array}{c} \alpha\left(1-F\left(q^{AI}_{l,1}(H)\right)\right) \\ +(1-\alpha)\left(1-F\left(q^*_{l,1}(L)\right)\right) \end{array}\right]}$$

The next step is to determine how does such a redistribution alternate the welfare of type L workers, that is, we should find the value of $\triangle^{AI}_{L,gain,31} - \triangle^{AI}_{L,loss,31}$. We have

$$\begin{aligned} & \triangle^{AI}_{L,gain,31} - \triangle^{AI}_{L,loss,31} \\ = & \left(1-F\left(q^*_{l,1}(L)\right)\right) \\ & \times\left[\bar{k}_{31} - \beta p_L\left(1-F\left(q^*_{s,1}\right)\right)\left(\beta p_L\left(1-F\left(q^*_{m,1}(L)\right)\right)+\beta\left(1-p_L\right)\frac{1}{\left(1-F\left(q^*_{l,1}(L)\right)\right)}\right)\right] \end{aligned}$$

Rearranging, we obtain

$$\begin{aligned} & \text{Sign}\left\{\triangle^{AI}_{L,gain,31} - \triangle^{AI}_{L,loss,31}\right\} \\ = & \text{Sign}\left\{\begin{array}{cc} & \beta p_H\left(1-F\left(q^*_{s,1}\right)\right)\left\{\beta p_H\left(1-F\left(q^{AI}_{m,1}(H)\right)\right)+\beta\left(1-p_H\right)\frac{1}{\left(1-F\left(q^{AI}_{l,1}(H)\right)\right)}\right\} \\ - & \beta p_L\left(1-F\left(q^*_{s,1}\right)\right)\left(\beta p_L\left(1-F\left(q^*_{m,1}(L)\right)\right)+\beta\left(1-p_L\right)\frac{1}{\left(1-F\left(q^*_{l,1}(L)\right)\right)}\right) \end{array}\right\} \end{aligned}$$

Under the condition $p\left(1-F\left(q_{m,1}^{*}(p)\right)\right)$ being increasing in p, we have $p_H\left(1-F\left(q_{m,1}^{*}(p_H)\right)\right) > p_L\left(1-F\left(q_{m,1}^{*}(p_L)\right)\right)$. According to Bi and Langot (2015), $q_{m,1}^{AI}(H) > q_{m,1}^{*}(H)$, so that $\beta p_H\left(1-F\left(q_{m,1}^{AI}(H)\right)\right) > \beta p_H\left(1-F\left(q_{m,1}^{*}(H)\right)\right)$ as $F(.)$ is a decreasing function. So we have $\beta p_H\left(1-F\left(q_{m,1}^{AI}(H)\right)\right) > \beta p_L\left(1-F\left(q_{m,1}^{*}(L)\right)\right)$. However, the sign of the expression is still ambiguous because of the $(1-p)p$ effect: the comparison between $(1-p_H)\frac{1}{\left(1-F\left(q_{l,1}^{AI}(H)\right)\right)}$ and $(1-p_L)\frac{1}{\left(1-F\left(q_{l,1}^{*}(L)\right)\right)}$ is ambiguous. On one hand $(1-p_H) < (1-p_L)$, on the other hand, $\frac{1}{\left(1-F\left(q_{l,1}^{AI}(H)\right)\right)} > \frac{1}{\left(1-F\left(q_{l,1}^{*}(L)\right)\right)}$. The $(1-p)p$ effect is related to the terms with $(1-p_t)$ in the expression. So that the presence of this $(1-p)p$ effect could again either make cross-subsidization for type L better or worse.

It is straightforward to find that the aggregate welfare improvement is measured by

$$\begin{aligned} dW_{31} &= \alpha\times\left(\triangle_{H,gain,31}^{AI}-\triangle_{H,loss,31}^{AI}\right)+(1-\alpha)\times\left(\triangle_{L,gain,31}^{AI}-\triangle_{L,loss,31}^{AI}\right) \\ &= \alpha\times\delta_l\times\left\{\begin{array}{c}\left[\left(1-F\left(q_{l,1}^{*}(L)\right)\right)-\left(1-F\left(q_{l,1}^{AI}(H)\right)\right)\right] \\ \times\left[\bar{k}_{31}-\beta p_L\times\left[\left(1-F\left(q_{m,1}^{*}(L)\right)\right)\beta p_L\left(1-F\left(q_{s,1}^{*}\right)\right)\right]\right]\end{array}\right\}\times u'(z)\,dz \end{aligned}$$

According to Bi and Langot (2015), we have $\left(1-F\left(q_{l,1}^{*}(L)\right)\right)-\left(1-F\left(q_{l,1}^{AI}(H)\right)\right) > 0$, so

$$\begin{aligned} &\operatorname{Sign}\{dW_{31}\} \\ =\ &\operatorname{Sign}\left\{\left[\bar{k}_{31}-\beta p_L\times\left[\beta p_L\left(1-F\left(q_{m,1}^{*}(L)\right)\right)\left(1-F\left(q_{s,1}^{*}\right)\right)\right]\right]\right\} \end{aligned}$$

When the effect of $(1-p)p$ is absent, it could be shown that $\operatorname{Sign}\{dW_{31}\} = \operatorname{Sign}\left\{\begin{array}{c}\beta^2 p_H^2\left(1-F\left(q_{m,1}^{AI}(H)\right)\right) \\ -\beta^2 p_L^2\left(1-F\left(q_{m,1}^{*}(L)\right)\right)\end{array}\right.$ which is always positive under our condition that $p\left(1-F\left(q_{m,1}^{*}(p)\right)\right)$ being increasing in p. When the effect of $(1-p)p$ is present, the value of $\bar{k}_{31}$ is again larger.

We now turn to Situation 3: Front-loading unemployment benefit from the second period to the first period. Define $\bar{k}_{21}$ such that every unit of third period unemployment benefit is equivalently redistributed as $\bar{k}_{21}$ unit to the second period unemployed. We have

$$\bar{k}_{21} = \frac{\alpha\left(1-F\left(q_{l,1}^{AI}(H)\right)\right)\times\beta p_H\times\left(1-F\left(q_{m,1}^{AI}(H)\right)\right)+(1-\alpha)\left(1-F\left(q_{l,1}^{*}(L)\right)\right)\times\beta p_L\times\left(1-F\left(q_{m,1}^{*}(L)\right)\right)}{\left[\alpha\left(1-F\left(q_{l,1}^{AI}(H)\right)\right)+(1-\alpha)\left(1-F\left(q_{l,1}^{*}(L)\right)\right)\right]}$$

The next step is to determine how does such a redistribution of a small magnitude dz alternate the welfare of type L workers, that is, we should find the value of $\triangle_{L,gain,21}^{AI} - \triangle_{L,loss,21}^{AI}$. We have

$$\begin{aligned}\triangle^{AI}_{L,gain,21}-\triangle^{AI}_{L,loss,21} &= \left(1-F\left(q^{*}_{l,1}(L)\right)\right)\times\bar{k}_{21}\times u^{'}(z)\,dz \\ &\quad -\beta p_L\times\left(1-F\left(q^{*}_{l,1}(L)\right)\right)\left(1-F\left(q^{*}_{m,1}(L)\right)\right)\times 1\times u^{'}(z)\,dz \\ &= \left(1-F\left(q^{*}_{l,1}(L)\right)\right)\left[\bar{k}_{21}-\beta p_L\left(1-F\left(q^{*}_{m,1}(L)\right)\right)\right]\end{aligned}$$

So that

$$\begin{aligned}&\mathrm{Sign}\left\{\triangle^{AI}_{L,gain,21}-\triangle^{AI}_{L,loss,21}\right\} \\ =\ &\mathrm{Sign}\left\{\left[\bar{k}_{21}-\beta p_L\left(1-F\left(q^{*}_{m,1}(L)\right)\right)\right]\right\} \\ =\ &\mathrm{Sign}\left\{\beta p_H\left(1-F\left(q^{AI}_{m,1}(H)\right)\right)-\beta p_L\left(1-F\left(q^{*}_{m,1}(L)\right)\right)\right\}\end{aligned}$$

Under our condition, it is always greater than zero. The $(1-p)\,p$ effect is absent here. The aggregate welfare is measured by

$$dW_{21}=\alpha\times\left(\triangle^{AI}_{H,gain,21}-\triangle^{AI}_{H,loss,21}\right)+(1-\alpha)\times\left(\triangle^{AI}_{L,gain,21}-\triangle^{AI}_{L,loss,21}\right)$$

After calculation, we have

$$\begin{aligned}dW_{21} &= \alpha\times\left(\triangle^{AI}_{H,gain,21}-\triangle^{AI}_{H,loss,21}\right)+(1-\alpha)\times\left(\triangle^{AI}_{L,gain,21}-\triangle^{AI}_{L,loss,21}\right) \\ &= \alpha\delta_l\times\left[\begin{array}{c}\left[\left(1-F\left(q^{*}_{l,1}(L)\right)\right)-\left(1-F\left(q^{AI}_{l,1}(H)\right)\right)\right] \\ \times\left[\bar{k}_{21}-\beta p_L\times\left(1-F\left(q^{*}_{m,1}(L)\right)\right)\right]\end{array}\right]\times u^{'}(z)\,dz\end{aligned}$$

And so

$$\begin{aligned}&\mathrm{Sign}\{dW_{21}\} \\ =\ &\mathrm{Sign}\left\{\left[\bar{k}_{21}-\beta p_L\times\left(1-F\left(q^{*}_{m,1}(L)\right)\right)\right]\right\} \\ =\ &\mathrm{Sign}\left\{\beta p_H\left(1-F\left(q^{AI}_{m,1}(H)\right)\right)-\beta p_L\left(1-F\left(q^{*}_{m,1}(L)\right)\right)\right\}\end{aligned}$$

which is always positive under our condition. **Q.E.D.**

Chapter 3

Holdup and Hiring Discrimination with Search Friction

Coauthored with Ms. Yuanyuan LI

Abstract

A holdup problem on workers' skill investment can arise when employers adopt discriminatory hiring norm to extract higher than socially optimal profit. When hiring priority is determined by both productivity-dependent (skill level) and -independent characteristics (discrimination), skill investment decision becomes strategic between the discriminated and favored group. We consider frictional markets with either posted or bargained wage (fixed sharing rule). With posted wage, depending on market tightness there may be equilibrium or multiple equilibria on skill investment. With discriminatory hiring, if in equilibrium both groups stay high skilled, both are worse off and firms better off; In any equilibrium where one group underinvest, the other group remain high skilled and are better off, while firms are worse off with discrimination. With bargained wage, similar equilibrium where the favored group underinvest exists, and firms incur cost for an intermediate range of bargaining power when they discriminate.

Keywords: Discrimination; Directed Search; Pre-matching Investment

JEL classification: J7, J42

3.1 Introduction

A holdup problem arises when some investment is sunk ex ante by one party, and the payoff is shared with that one party's trading partner. Since cost has no other use once sunk, that trading partner will have every incentive to squeeze the profit at the ex post stage. In an important study on such a problem in a labor market with search friction, Acemoglu and Shimer (1999b) show that with firms' sinking capital and ex post wage bargaining, the equilibrium is always inefficient, since wages paid ex post can be so high such that firms' ex ante incentive of investment is harmed; while if firms are able to post wages to direct workers' search, then the holdup problem to firms' investment no longer appears; the efficiency can be achieved, because wage posting allows workers to observe offers and choose where to apply, and it induces workers to optimize their expected payoff from application by making trade-off between every wage they observe and the probability of obtaining it. Within conventional wage posting framework, we spot another source of inefficiency in a holdup problem where workers sink skill investment cost: when the market is crowded for the firms, by adopting a discriminatory hiring norm firms are able to expropriate higher than socially optimal level of profit, and this has the consequence of discouraging the investment incentives for both the favored and discriminated groups. We analyze the impact of such rent seeking behavior of firms on the structure of market segmentation, and on the workers' skill investment incentives.

When discrimination is absent, the wage posting economy with workers' ex ante skill investment attains efficiency in the equilibria, and we show which equilibrium emerges depends on the rivalry between the log return to skill and the market tightness (workers/firms ratio) which measures the degree of market competition. The fundamental reason behind this efficiency result is that skill achievement is a quality which can be legally written into the wage contracts. It is a different story when other (binary) characteristics which are not closely related to productivity, such as gender, race, height, origin etc. enter also into firms' preference. Under equal pay legislation, posted wages can not be conditioned explicitly on these characteristics; however, if firms still select workers according to their preference on these characteristics, a separating equilibrium can result where separate firms post different levels of wages, and workers of different groups sort themselves and apply to different wages: the market is then endogenously segregated. On the side of firms, they have incentive to adopt such discriminatory hiring norm, when workers' return to skill investment is sufficiently high; in that case discrimination allows them to grasp higher than the socially optimal level of operating profit. On the side of the workers, it proves that both the discriminated group and favored group are worse off: for the former, it is because discrimination discretely reduces the labor market opportunity of these workers, who anticipate discrimination, then demand lower wages, which makes them cheaper to hire; for the latter, it is so because when firms are able to hire the discriminated workers cheaply, it is as if firms enjoy larger "market power", which allows them to suppress further the undiscriminated workers' expected payoff. Naturally, anticipating discrimination, all groups expect lower payoff from search, jeopardizing their skill investment incentives.

A key feature of our study is the multidimensionality of characteristics based on which workers are ranked. On one hand, there is ranking by productivity-dependent type identity: workers are either high skilled (type H) or low skilled (type L); high skilled have priority to low skilled simply because such ranking gives firms higher profit. On the other hand, there is ranking by productivity-independent group identity: workers belong either to the favored (group a) or the discriminated group (group b). The resulting ranking schedule has the following order: $aH \succ bH \succ aL \succ bL$. It reads: given any skill level, group a are preferred to group b; the high skilled are always preferred to low skilled. Under such an "intertwined" ranking order, the skill investment decision for different groups becomes strategically interdependent. Focusing on Nash pure strategy equilibrium on skill investment, in the wage posting economy, we find that depending on the value of market tightness there can be equilibrium or multiple equilibria on skill investment due to that interdependence. Compared to the case without discrimination, when the market is very crowded (market tightness is small) for the firms, discrimination is profitable for firms and all the workers are worse off; as the tightness further increases, both group can choose low skill and in equilibrium whenever one group underinvest, the other group remain high skilled and are better off, while the firms are worse off with discrimination. In particular, the equilibrium where the favored group underinvest, while the discriminated group choose to remain high skilled exists; And in this case firms' profits drop since workers' underinvestment in skill leads to lower average productivity in the economy compared to the case where discrimination is absent.

In the economy where wages are bargained (determined according to a fixed sharing rule) after matching hence do not direct search, we find similar equilibrium where the favored group underinvest, hence earn lower expected payoff compared to the case without discrimination within a certain region of bargaining power; in such an equilibrium, surplus is transferred from firms and favored group to discriminated group. Firms' profits are piecewise monotone, because increase of workers' bargaining power can increase workers' incentive of skill investment, hence discretely improves the market skill composition and average productivity. We also find that there is an intermediate range of workers' bargaining power for values of which firms are worse off by discriminating, due to discouraged skill investment from discriminated group. All in all, the key difference between wage posting and wage bargaining is that the actual wage now exogenously pegs on the productivity, and firms can no longer manipulate their market power by translating their discriminatory preference into constantly lower wages.

3.1.1 Relation to the literature

Job search process is an important channel through which discrimination keeps functioning in the labor market. Several papers have highlighted the impact of discrimination through job search channel to the wages gaps. To name a few, Pendakur and Woodcock (2010) show that the existent glass ceilings for the immigrant and minority workers may be attributed by large measure to their poor access to the jobs in high-wage firms; As well, in an important article from Ritter and Taylor (2011), they show that most of the disparity in unemployment rate could not be explained by cognitive skills that emerge at an early stage, although for wage gap it could be the case. This result concerning the unemployment disparity is confirmed by the finding that this disparity is still significant even for workers of similar skill levels.

Our work is most closely related to the directed search literature[1]. In this literature, search frictions are derived endogenously through agents' sequential strategic interactions. Taking into account strategic interaction allows search externality to be internalized. The resulting economy remains competitive, albeit with a non-Walrasian market structure, and prices play an allocative role to achieve efficiency. To the best of our knowledge, among the discrimination literature with search friction, only two of them are built upon wage posting context. Lang, Manove, and Dickens (2005, hereafter LMD) show that a discriminatory hiring rule could lead to labor market segmentation and significant wage gap with even a negligible difference in productivity; however, the discriminated group turn out to have lower unemployment rate, which is in sharp contrast with evidence. Merlino (2012) aims at improving the result of LMD (2005). He considers further the pre-matching investment from the firms' side, and obtain technology dispersion and realistic unemployment gap. His results rely on the strong assumption that there is more discrimination in the high technology sector, and he is silent on the workers' skill levels. Our paper differs from theirs, in that our focus is to analyze how hiring discrimination could distort workers' skill investment incentives and the structure of market segmentation.

[1]This literature is sometimes also termed as wage posting game with coordination friction.

While the setup of wage bargaining (no information of level of wage before matching) is more prevalent, it neglects an important trade-off that the workers make to some extent in their search for jobs: the wage and the probability of obtaining it. This endogenous link between wage and employment probability is especially important, since wages convey information on whether the employers discriminate. Having information of wages available before matching, workers are able to adjust accordingly their search strategy to avoid being discriminated. Workers apply to certain wage only when their expected payoff (wage times the employment probability) from this application attains certain level, and a high wage which attracts also the favored group discretely lowers the probability of employment for the discriminated group to such an extent that the expected payoff for the latter at these high wage firms does not meet the expected market payoff. This setup is supported by Lang and Lehmann (2012) and Heckman (1998), who mention that workers do not apply randomly and they actually avoid prejudiced employers to some extent, which implies between-group search externality is taken into account by the discriminated workers. Moreover, it is well known that within-group search externality may be prevalent when wages are bargained; while in wage posting context, we are able to abstract from search externality and focus on discrimination. Hall and Krueger (2010) use U.S. data to show that fraction of posted and bargained wages are both around one third. They also document a negative relationship between the education level and precise information concerning the expected pay. Brenzel, Gartner and Shnabel (2013) focus on the employer's side of the study in Germany, and showed that around two thirds of the wages are posted, and the bargained wages are more likely set for those with higher education and qualification. The message is that not only is wage posting a prevalent wage determination process in the labor market, more importantly, it is also dominant in the relatively low skilled sector.[2] Within our context, employers can not post wages contingent on workers' group identity which is irrelevant to productivity, which could be understood as due to the functioning of the equal opportunity legislation.

Literature addressing discrimination problem in random search context is vaster. However, to have tractable such model convenient for linking to evidence, the introduced discrimination is usually taste-based, hence to obtain realistic outcome may often require making compromise on assuming ex ante differences in parameters governing relevant characteristics. Rosen (1997) is an exception and shows that discrimination can result even if there are no differences across groups. Job opportunities arrive stochastically, minority workers choose reservation productivities above which they accept the job; To avoid majority workers who

[2]It is consistent with our knowledge that the more skilled workers, whose number is comparatively small, usually receive more attention and protections.

are always preferred, they choose to accept jobs even with low reservation wages. Although private information is the key element in Rosen (1997)'s model, search externality remains the main channel for the functioning of the discrimination mechanism. Our focus is on how the ranking order of firms contributes to strategic interdependence in workers' skill investment decisions, and search externality is internalized when search is directed.

There is also the important statistical discrimination literature[3] which emphasizes the role of asymmetric information on qualities related to the productivity. One strand of this literature derives group inequalities endogenously even in the absence of ex ante group difference on relevant characteristics. Their mechanism is that decision makers' asymmetric beliefs on relevant characteristics of members for different groups could subsequently dim unfavored agents' incentive on investment on payoff-relevant technology, which in turn justifies the firms initial beliefs. Our context is different from this literature mainly in the point that, instead of relying on the information friction which plays central role in generating the pessimistic outcome, we work through a sequential game where agents could correctly anticipate the pessimistic outcomes, hence choose to react accordingly in a rational way.

The paper is organized as follows. Section 2 analyzes the case without discrimination. We then move to the economy with discrimination in Section 3. In section 4, we consider the case of wage bargaining. Section 5 discusses, and then we conclude.

3.2 The model without discrimination

We start with a context without hiring discrimination. Consider an economy populated by two kinds of agents, the workers and the firms. The number of workers is N,[4] with the index $i \in \{1,2,...,N\}$, and the number of firms is M, with the index $j \in \{1,2,...,M\}$. Define the market tightness as $\beta \equiv \frac{N}{M}$.

We introduce a pre-matching investment stage in a standard wage posting game. Each job seeker makes a skill investment decision before entering into the labor market. This skill choice is assumed to be binary, such that if the worker decides to become highly skilled, an investment cost E_H is paid, and otherwise E_L, with $E_H > E_L$. A highly skilled job seeker who pays E_H is capable of producing y_H; while a low skilled could only deliver y_L. It would be

[3] We refer readers to the survey from Fang & Moro (2010).

[4] As noted by Lang, Manove, and Dickens (2005), the number N could be regarded as the expected number of entrants (job seekers) from the firms' perspective.

useful to understand the formulation in the following way: workers who enter labor market after a longer period of training at school expect to receive higher expected income compared to those who spend a shorter period in schooling and enter the market at an earlier stage; the opportunity cost for the former is E_H, and for the latter is E_L. We assume that workers' skill level is public information. The costs $\{E_L, E_H\}$ and productivities $\{y_L, y_H\}$ are exogenous, but should satisfy some conditions which will be specified later.

Firms are ex ante identical. Having observed the distribution of skill attainment of job seekers, they post wages conditional on skills. If firms choose to attract a high skilled worker, they post w_H, and the surplus after matching is $y_H - w_H$; in case a low-skilled worker is searched for, w_L is announced and the surplus is $y_L - w_L$. We emphasize that skill is a characteristic of workers which the wage contracts can be conditioned on; this is in sharp contrast to other qualities such as gender, race, height etc. which, under equal pay legislation, should not be conditioned on; so when firms distinguish workers according to these latter qualities, the wage contract becomes "incomplete";[5] by this, we will say that firms discriminate.

The timing follows that of a standard wage posting game, augmented with a pre-matching skill investment stage (Stage 0):

Stage 0: *Workers choose skill level, and pay either E_H or E_L.*

Stage 1: *Firms observe skills of job seekers, and announce the wage (w_L, w_H).*

Stage 2: *Workers observe the wages offers, and choose which wage to apply to.*

Stage 3: *Firms select workers from the received applications, and they select workers with same skill levels with equal probability. Then the production is carried on, and payoffs are realized.*

We will focus on subgame perfect equilibria. Firms choose wages to maximize profits, and workers choose firstly the skill level and then which wage to apply to, in order to maximize the expected payoff from search.

[5] Incompleteness of contract is the source of inefficiency for the holdup problem. See Acemoglu and Shimer (1999) for related literature.

3.2.1 Specification of the strategies, matching probabilities, and payoff functions.

To write agents' payoffs, it is a routine procedure in the directed search literature to first derive the matching functions. This section provides a quick summary for the general understanding of the context.

Define a type-t job seeker i's strategy as a vector of probabilities $\Theta_t^i = \left(\theta_t^{i1}, ..., \theta_t^{iM}\right)$, where θ_t^{ij} is the probability with which the type-t worker i applies to firm j, and $t \in \{L, H\}$. It holds that $\sum_j \theta_t^{ij} = 1$ for any i and t. As in the literature, it is convenient[6] to proceed with a transformation of variable. We define q, as *expected number of applications received per firm;* it is also called the expected queue length.

Denote q^j as the queue length of firm j, and q_t^j as the queue length of the type-t workers in firm j. If a firm attracts both high and low skilled, we have $q^j = q_L^j + q_H^j$, where q_L^j and q_H^j are the queue length of the corresponding workers in firm j. Since we only consider symmetric equilibria, for a given firm j, θ_t^{ij} has the same value for any type t job seeker, so we denote $\theta_L^{ij} = \theta_L^j$ and $\theta_H^{ij} = \theta_H^j$ for any j. By definition, q_t^j is the number of workers of type t in firm j times their application probability: $q_H^j = N_H \times \theta_H^j$, and $q_L^j = N_L \times \theta_L^j$ for any j, where N_L and N_H are the total number of low skilled and high skilled workers respectively.

Firms. A particular firm j matches with a worker if after the search stage at least one worker appears, which happens with probability $1 - \left(1 - \theta_t^j\right)^{N_t}$. The probability that no job seeker sends application to this firm is $\left(1 - \theta_t^j\right)^{N_t}$, and $1 - \left(1 - \theta_t^j\right)^{N_t}$ is the probability of receiving at least one application from type t workers. According to the aboved defined relationship $q_t^j = N_t \theta_t^j$, the probability $\left(1 - \left(1 - \theta_t^j\right)^{N_t}\right)$, goes to $\left(1 - e^{-q_t^j}\right)$ when $N_t \to \infty$. This probability is increasing in q, which means that the more the expected number of applicants, the higher the probability that the firm could fill the vacancy. The firm chooses wage to maximize their expected profit, which is the product of the probability of meeting a worker of this type and the net surplus, $\left(1 - e^{-q_t^j}\right) \times \left(y_t - w_t^j\right)$, where $y_t \in \{H, L\}$. In the following we may occasionally suppress the superscript j whenever it does not raise confusion.

[6]When the number of firms and workers are large, it is no longer convenient to operate with the workers' application strategy θ_i^j, because it will tend to zero in the symmetric mixed strategy equilibrium.

As shown by Shi (2006), in case workers of both skill levels appear in the market, it is optimal for firms to post both w_L and w_H to attract both skill types. Furthermore, firms rank the high skilled in priority to the low skilled, that is, firms will only consider hiring the low skilled workers when they did not receive any application from high skilled workers, an event which happens with probability e^{-q_H}. Then the expected profit *(from attracting both types of workers)* is

$$(1-e^{-q_H})\times(y_H-w_H)+e^{-q_H}(1-e^{-q_L})\times(y_L-w_L)$$

Job seekers. Job seekers observe all the wages **w** announced by firms, and choose which wage to apply to. Consider a particular job seeker. Conditional on visiting a particular firm j, his probability of employment in that firm is $\frac{1-\left(1-\theta_H^j\right)^{N_H}}{N_H\theta_H^j}$ if he is high skilled, and $\left(1-\theta_H^j\right)^{N_H}\times\frac{1-\left(1-\theta_L^j\right)^{N_L}}{N_L\theta_L^j}$ if low skilled (see Appendix for more details). And these probabilities become $\frac{1-e^{-q_H^j}}{q_H^j}$ and $e^{-q_H^j}\frac{1-e^{-q_L^j}}{q_L^j}$ when $N\to\infty$ and $M\to\infty$. Notice that $\frac{1-e^{-q_t}}{q_t}$ is decreasing in q_t: the higher the expected number of applicants in this firm competing this job with him, the lower the probability with which this job seeker will be employed. Also notice that the employment probability of the low skilled workers is a product of e^{-q_H} and $\frac{1-e^{-q_L}}{q_L}$, where the former governs the between-group competition effect, and $\frac{1-e^{-q_L}}{q_L}$ governs the within-group competition effect.

We remark that since q is a function of job seekers' application strategy, it depends on w. We now look more closely into their causal relationship. We should distinguish two terms: (1) each job seeker's *expected payoff from application*, and (2) her *expected "market" payoff*. The expected payoff from application is the payoff that a worker receives when applying to a certain firm, namely, a product of the wage and the probability of obtaining it at that firm, namely, $\frac{1-e^{-q_H}}{q_H}\times w_H$ for the high skilled, and $e^{-q_H}\frac{1-e^{-q_L}}{q_L}\times w_L$ for the low skilled. The expected market payoff, denoted by U_t, is the maximum level of the expected payoff from application in the equilibrium. U_t is regarded as an aggregate variable, which is assumed to be invariant with respect to any variation of an individual agent's strategy. Consider a particular type H job seeker. He is willing to send application to a particular firm j, if and only if his expected payoff $\frac{1-e^{-q_H^j}}{q_H^j}\times w_H^j$ from doing so is equal or greater than the expected market payoff U_H. By the definition that U_H is the maximum level attainable, we have

$$q_H^j\begin{cases}>0 & \text{if } \frac{1-e^{-q_H^j}}{q_H^j}\times w_H^j=U_H\\ =0 & \text{if } \frac{1-e^{-q_H^j}}{q_H^j}\times w_H^j<U_H\end{cases}$$

Job seekers make trade-off between the wage and the probability obtaining it. To highlight the dependence of q_t on w_t, we could rewrite the above expressions as

$$q_H^j \begin{cases} > 0 & \text{if } w_H^j > U_H \\ = 0 & \text{if } w_H^j \leq U_H \end{cases}$$

Indeed, the employment probability $\frac{1-e^{-q_H}}{q_H}$ is a number which belongs to the interval $(0,1)$; if the wage is too low such that $w_H^j \leq U_H$, it will attract no workers, q_H is zero; if $w_H^j > U_H$, there is always a positive value q_H which satisfies $\frac{1-e^{-q_H^j}}{q_H^j} \times w_H^j = U_H$.

We now formalize the notion of equilibrium and will proceed in two steps. In the first, we state the notion of equilibrium for the wage posting subgame given a skill distribution of workers. In the second, we distinguish two notions on how workers optimally choose their skill level. For the wage posting subgame, the solution concept will be standard subgame-perfect competitive equilibrium (SPCE) similar to LMD (2005). As emphasized by LMD (2005), it is *'a simplification of standard subgame-perfection in which the aggregate variables are assumed constant with respect to the changes in the strategy of an individual agent.'* Given the fraction of $\alpha \in [0,1]$ of high skilled workers, an equilibrium, symmetric among workers, consists of the expected market payoffs (U_H, U_L), each firms' strategy (w_H^*, w_L^*), and workers' strategies (q_H^*, q_L^*), that satisfy the following requirements: (i) each firm's posted wage (w_H^*, w_L^*) is a best response to the other firms' strategy and to the workers' strategies (q_H^*, q_L^*), on the assumption that the market expected payoff ($U_H(\mathrm{w})$, $U_L(\mathrm{w})$)[7] remains fixed at $U_H(\mathrm{w}^*)$ and $U_L(\mathrm{w}^*)$, and is invariant to the firm's own wage; and (ii) $q_t^*(\mathrm{w})$ with $t \in \{L,H\}$ is a best response of each worker to any wage vector w, and to the choice of $q_t^*(\mathrm{w})$ by all other workers.

Now we turn to workers' skill decision which determines the value of α. Workers choose to be high skilled whenever $U_H - E_H > U_L - E_L$, then the fraction of high skilled workers is such that $\alpha = 1$. Workers choose to be low skilled whenever $U_H - E_H < U_L - E_L$, then the fraction of high skilled workers is such that $\alpha = 0$. Workers are indifferent between high and low skill when $U_H - E_H = U_L - E_L$. In the case with indifference, we should make distinctions on two notions specified as follows:

- Notion 1 (perfect correlation): Whenever indifferent, all workers (within a group) choose either high or low skilled.

[7]"w" denotes the wage vector of all the firms.

- Notion 2 (no perfect correlation): Whenever indifferent, α represents individual worker's probability of choosing to be high skilled.

Under Notion 1, all workers in one group, whenever indifferent between two alternatives (L or H), will randomize towards the same direction: that is to say, we consider the group of workers as a whole, or there is perfect correlation on their skill choices; as a result, α does not represent an individual's probability of choosing high skilled and takes either the value 0 or 1, $\alpha \in \{0,1\}$. Under Notion 2, each worker can randomize between high and low skill whenever they are indifferent; since each worker can end up either high or low skilled, after this randomization, high and low skilled workers can be present at the same time, in contrast to Notion 1; The equilibrium value of α, denoted by α^*, will be endogenously determined by the indifference equation, in which case $\alpha^* \in (0,1)$.

In the following analysis in this section without discrimination, we will derive equilibrium under both notions because technically Notion 1 could be regarded as a subproblem of Notion 2. However, we will focus exclusively on the study under Notion 1 when we turn to the section with discrimination. By doing so, we are able to abstract from some equilibrium which only arises under theoretical rigor but at the same time yields insights to a limited extent and induces unnecessary complexity in analysis.

Solution of the wage posting subgame. According to Burdett, Shi and Wright (2001), the definition of equilibrium is equivalent to a problem where firms choose wages to maximize their expected profit, taking into account the best responses of other firms as well as of the job seekers. As we consider a large economy, in which a single firm's deviation does not alter the expected market payoff U_t, the market payoff can be taken as given in the stage where a firm maximizes its profit, and will later be determined endogenously. Thus, we consider the optimization problem in which a (deviating) firm j chooses w_t with $t \in \{L,H\}$ to maximize its expected profit, taking expected market payoff U_t (other firms' responses) and the functional relationship between w_t and q_t (job seekers' responses) as given. When the firm attracts a single skill type, the equilibrium can be solved by[8]

$$\begin{array}{ll} \max_{w_t} & (1-e^{-q_t}) \times (y_t - w_t) \\ \text{s.to} & \frac{1-e^{-q_t}}{q_t} \times w_t = U_t \end{array}$$

for $t \in \{L,H\}$. For a given U_t, solving w_t from the constraint, substituting it into the objective function, and maximizing with respect to q_t, we can obtain an optimal functional relationship

[8]The program could also be understood as a deviating firm's profit maximization, taking all the other firms' best response as given. The other firms' best response is w_t^*, which satisfies $\frac{1-e^{-q_t^*}}{q_t^*} w_t^* = U_t$.

between q_t^* and U_t. Using this obtained relationship, with the help of the constraint, we then achieve an optimal functional relationship between w_t^* and q_t^*. And since in symmetric equilibrium all firms will post the same wages, so that all workers will apply to each firm with equal probability, by definition of q we have $q_H^j = q_H^* = \frac{N}{M} = \beta$ and $q_L^j = q_L^* = \frac{N}{M} = \beta$, when either all workers are high skilled or all workers are low skilled. This program is applicable under both Notion 1 and relevant situations of Notion 2. Under Notion 2, in addition, when workers are indifferent, both skills can be present at the same time, then firms attract both skill types as in Shi (2006), the problem is

$$\begin{array}{ll} \max_{w_H,w_L} & (1-e^{-q_H}) \times (y_H - w_H) + e^{-q_L}(1-e^{-q_H}) \times (y_L - w_L) \\ \text{s.to} & \frac{1-e^{-q_H}}{q_H} \times w_H = U_H \\ & e^{-q_H}\frac{1-e^{-q_L}}{q_L} \times w_L = U_L \end{array}$$

Solving the program, we could obtain $U_H = e^{-q_H^*}(y_H - y_L) + e^{-q_H^* - q_L^*} y_L$, while $U_L = e^{-q_H^* - q_L^*} y_L$; at the equilibrium, we have $q_H^* = \frac{N_H}{M}$, $q_L = \frac{N_L}{M}$ and $q_H^* + q_L^* = \beta$. At last, it is important to remark that q_t depends on w_t continuously, as remarked by Shi (2002). In this way, a marginal change of wage w_t can only lead to a marginal modification on the expected number of applicants q_t. By the definition of q, under Notion 2, we have $q_H^*(\alpha) = \frac{N_H}{M} = \alpha\beta$, and $q_L^*(\alpha) = \frac{N_L}{M} = (1-\alpha)\beta$. Once the skill investment choice pins down the value of α^*, we obtain $q_H^* = \alpha^*\beta$, and $q_L^* = (1-\alpha^*)\beta$. In the next section, we establish the decentralized market equilibrium and examine its properties. It is convenient to start with Notion 2.

Decentralized Market Equilibrium without discrimination under Notion 2

Firms' wage offers are conditioned on job seekers' skill levels, so we first consider the skill investment decision of job seekers at first stage. Denote α as the fraction of the job seekers who choose to invest in high skill, so the remaining fraction $(1-\alpha)$ is low skilled. Under Notion 2, α is also the probability with which a job seeker chooses to invest in high skill, by virtue of the Law of Large Number. Let α^* denote the equilibrium fraction of high skilled job seekers on the total population. There are three cases:

Case (1). $\alpha^* = 1$. *All job seekers invest in high skill.*

Case (2). $\alpha^* \in (0,1)$. *Some invest in high skill, while the remaining in low skill.*

Case (3). $\alpha^* = 0$. *All invest in low skill.*

With Case (1) and Case (3), there exists only one skill level in the market, and since skills can be conditioned on wages, there is only one wage posted in equilibrium. However, the market with Case (2) features two skill levels. Shi (2006) establishes that in this case it is optimal for firms to attract both skill types, while ranking the high skilled in priority to the low skilled. We now show that the rivalry between the market competition (captured by market tightness β) and the magnitude of the return to skill ratio $\frac{y_H - y_L}{E_H - E_L}$ are crucial in the determination of which of the above three cases may prevail.

Proposition 1. (return to skills) *Given the return to skill ratio $\frac{y_H - y_L}{E_H - E_L}$, define $\hat{\beta}$ as $\frac{y_H - y_L}{E_H - E_L} = e^{\hat{\beta}}$.*

(i) when $0 < \beta \leq \hat{\beta}$, the unique equilibrium is such that all job seekers choose high skill, i.e. $\alpha^* = 1$.

(ii) when $\beta > \hat{\beta}$, the equilibrium consists of a unique value $\alpha^* \in (0,1)$ which satisfies $\frac{y_H - y_L}{E_H - E_L} = e^{\alpha^* \beta}$.

(iii) when $\frac{y_H - y_L}{E_H - E_L} \leq 1$ such that $\frac{y_H - y_L}{E_H - E_L} = e^{\hat{\beta}}$ has no positive solution on $\hat{\beta}$, the unique equilibrium is $\alpha^* = 0$.

Proof. In the appendix.

When the value of return to skill $\frac{y_H - y_L}{E_H - E_L}$ is sufficiently large compared to e^{β}, which measures the intensity of competition of the market, job seekers find it a dominant strategy to invest in high skills; There is no incentive for them to deviate, and the output is highest among all the equilibria. When the value of $\frac{y_H - y_L}{E_H - E_L}$ is moderate, there exists an equilibrium where job seekers are indifferent from being high skilled or low skills; all firms find it optimal to attract both skill types; the output is lower compared to the previous equilibrium. At last, when the value of return to skill is sufficiently low, it does not provide them incentive to sink this fixed cost against the risky job search game they are going to play; the equilibrium level of output turns out to be the lowest.

Decentralized Market Equilibrium without discrimination under Notion 1

Under Notion 1, there is only one skill type present in the market. When firms attract a single skill type, they solve

$$\begin{array}{ll} \max_{w_t} & (1 - e^{-q_t}) \times (y_t - w_t) \\ \text{s.to} & \frac{1 - e^{-q_t}}{q_t} \times w_t = U_t \end{array}$$

In equilibrium, for the high skilled we have: $U_H = e^{-q_H^*} y_H$, $w_H^* = \frac{e^{-q_H^*} q_H^*}{1 - e^{-q_H^*}} y_H$, where $q_H^* = \beta$; and for the low skilled: $U_L = e^{-q_L^*} y_L$, $w_L^* = \frac{e^{-q_L^*} q_L^*}{1 - e^{-q_L^*}} y_L$, where $q_L^* = \beta$. So that we have

$$\alpha = \begin{cases} 1 & \text{when } e^{-\beta} y_H - E_H > e^{-\beta} y_L - E_L \\ 0 \text{ or } 1 & \text{when } e^{-\beta} y_H - E_H = e^{-\beta} y_L - E_L \\ 0 & \text{when } e^{-\beta} y_H - E_H < e^{-\beta} y_L - E_L \end{cases}$$

It turns out that the threshold which makes workers indifferent is the same as $\hat{\beta}$ established under Notion 2.

3.2.2 Constrained efficient allocations

The objective of this section is to find the efficient allocations in the centralized market, and evaluate whether the decentralized market attains its efficiency. The social planner maximizes the aggregate output, subject to the same matching friction as in the decentralized equilibrium. More precisely, the social planner chooses the fraction of workers to be high-skilled, divides firms into different groups to attract distinct compositions of workers, and assigns workers to match with a certain group of firms. With the same matching friction as before, the social planner is restricted to treat workers of the same skill level in the same way, and assures that workers of the same skill type must match with firms from the same group with the same probability.

We use the superscript "p" to label the equilibrium allocation chosen by the planner. Let α be the fraction of high skilled workers the social planner chooses, $\alpha \in [0,1]$. If the optimal arrangement is $\alpha^p = 1$, all job seekers are high skilled, and only one type of firms exists - those which attract high skilled workers. It is similar for $\alpha^p = 0$. If $\alpha^p \in (0,1)$, there are both high and low skilled job seekers and it is optimal for the planner to assign all firms to post wages for both the high and low skilled (shown in Shi (2006)). Furthermore, in the last case the planner can also manage the priority of firms' hiring workers with different skills - whether to prefer high skilled to low skilled or otherwise. Let R be the probability with which the firms rank high skilled workers in priority to the low skilled. And q_t is the expected number of applicants in a firm, $t \in \{L,H\}$, which governs how the planner assigns workers' applications. Thus, the social planner's problem is to maximize the following aggregate output

$$M \times \left[\begin{array}{c} (1-e^{-q_H})(R \times (1-e^{-q_L}) + e^{-q_L}) y_H \\ +(1-e^{-q_L})((1-R)(1-e^{-q_H}) + e^{-q_H}) y_L \end{array} \right]$$
$$-N \times (\alpha E_H + (1-\alpha) E_L)$$

If at least one high skilled visits a certain firm, with probability $(1-e^{-q_H})$, the firm hires this high skilled, either with probability 1 when no low skilled worker shows up, which happens with probability e^{-q_L}, or with probability R if there is at least one low skilled who shows up at the same firm, which occurs with probability $(1-e^{-q_L})$; it is similar for the case with low productivity. Since the firms and workers of the same skill are all identical from the planner's perspective, we have $q_H = \alpha\beta$ and $q_L = (1-\alpha)\beta$. The above objective includes all cases with different values of α. Solving the problem, we can see that the optimal ranking is that firms always prefer high skilled workers, i.e., $R^p = 1$, and we have the following proposition

Proposition 2. *The equilibrium labor allocation and skill investment choice are socially optimal.*

Proof. In the Appendix.

In the proof, we could also verify that the threshold $\hat{\beta}^p$ for skill investment coincides with $\hat{\beta}$ in the decentralized economy. That is, when $0 < \beta \leq \hat{\beta} = \log\frac{y_H - y_L}{E_H - E_L}$, it is socially optimal that workers all invest in high skill; when $\beta > \hat{\beta}$, it is socially optimal that a fraction α^* of workers invest in high, while the rest invest in low skill; and $\frac{y_H - y_L}{E_H - E_L} \leq 1$ such that $\hat{\beta}$ has no positive real solution, all invest in low skill. In the rest of the paper, we mainly focus on the first case, so that whenever workers are discouraged to underinvest, it is due to the effect of discrimination.

3.3 The model with hiring discrimination

We now introduce discrimination. Consider an economy where workers can be partitioned into two groups, group a and group b, according to certain trait which is irrelevant to productivity. Gender, for example, is such one possible binary partition of labor force. Denote the fraction of group a as γ, and the fraction of group b as $1-\gamma$. The two group of workers are ex ante identical in all other aspects. Discrimination modifies the matching functions of agents. Specifically, in order to formulate discrimination, we introduce a term x called hiring (ranking) rule specified by firms. To be precise, x could be understood as the probability with which the group a workers are selected when workers from both groups are present. The probability that a group a worker is employed by this firm is

$$F_a(q_a, q_b, x) = \frac{1-e^{-q_a}}{q_a} \times \left[x\left(1-e^{-q_b}\right) + e^{-q_b}\right]$$

Analogously, the probability that a group b worker is employed by this firm is

$$F_b(q_a,q_b,x)=\frac{1-e^{-q_b}}{q_b}\times\left[(1-x)\left(1-e^{-q_a}\right)+e^{-q_a}\right]$$

To understand these expressions, we have to notice that now when job seekers are considering their probability of being hired, they have to take into account of the impact from the competition with the other group. The parts $\frac{1-e^{-q_a}}{q_a}$ and $\frac{1-e^{-q_b}}{q_b}$ capture the within group competition, while the remaining parts with x capture the between group competition.

When $x=1$, firms hire group b workers only when none of the group a is present. Firms' preference is such that group a are preferred to group b, although both groups have identical productivity. This is what we call hiring discrimination. The employment probability for group a and group b workers become respectively,

$$F_a(q_a,q_b,1)=\tfrac{1-e^{-q_a}}{q_a}$$

$$F_b(q_a,q_b,1)=\tfrac{1-e^{-q_b}}{q_b}\times\left[e^{-q_a}\right]$$

Another interesting example is $x=\frac{1}{2}$. The employment probability for group a and group b workers becomes respectively

$$F_a\left(q_a,q_b,\frac{1}{2}\right)=\frac{1-e^{-q_a}}{q_a}\times\left[\frac{1}{2}\times\left(1-e^{-q_b}\right)+e^{-q_b}\right]$$

$$F_b\left(q_a,q_b,\frac{1}{2}\right)=\frac{1-e^{-q_b}}{q_b}\times\left[\frac{1}{2}\times\left(1-e^{-q_a}\right)+e^{-q_a}\right]$$

When $q_a=q_b$, we have $F_a\left(q_a,q_b,\frac{1}{2}\right)=F_b\left(q_a,q_b,\frac{1}{2}\right)$, both workers have identical employment probability, and it is as if there is no discrimination among workers. Furthermore, we have $F_a(q_a,q_b,1)\geq F_a\left(q_a,q_b,\frac{1}{2}\right)$ for any positive value of (q_a,q_b). Then the employment probability of workers from the preferred group (group a) is higher under discrimination than that in the case without discrimination. Similarly, when $F_b(q_a,q_b,1)\leq F_b\left(q_a,q_b,\frac{1}{2}\right)$ for any positive values of (q_a,q_b), i.e. the employment probability of the discriminated group (group b) is lower when there is hiring discrimination. In fact, the employment probability of group a (group b) is increasing (decreasing) in x.Thus, for a given positive pair of (q_a,q_b), there exists $\hat{x}\in(0,1)$ such that $F_a(q_a,q_b,\hat{x})=F_b(q_a,q_b,\hat{x})$, and $\hat{x}$ could be considered as the hiring rule without discrimination. Indeed x measures the intensity of firms' discriminatory preference. Given q_a and q_b, for $x\in[0,\hat{x})$, firms discriminate group a, and for $x\in(\hat{x},1]$ firms discriminate against group b. The closer x approaches to the extremes of the interval $[0,1]$, the more intensive the hiring discrimination is. In the rest of paper, we focus on the case $x=1$ such that group a achieve absolute priority to group b.

3.3.1 The case of strong discrimination: $x = 1$

Formally, we introduce two assumptions as in Merlino (2012). These assumptions help introduce some heterogeneity which is not related to productivity among the labor pool.

Assumption 1: Firms are not allowed to post wages which are dependent on the group identity.

Assumption 2: Firms prefer group a in the sense that firms only hire workers from group b when no group a workers are present, i.e. $x = 1$.

Same as the case without discrimination, workers visit a firm only when they can obtain the expect market payoff from applying to that firm. We denote in this section the expected market payoff of high skilled job seekers from group a and group b as U_{aH} and U_{bH} respectively. The above assumptions imply

$$U_{aH} = F_{aH}\left(q_{aH}, q_{bH}\right) \times w_H = \tfrac{1-e^{-q_{aH}}}{q_{aH}} \times w_H$$

$$U_{bH} = F_{bH}\left(q_{aH}, q_{bH}\right) \times w_H = \tfrac{1-e^{-q_{bH}}}{q_{bH}} e^{-q_{aH}} \times w_H$$

For $\beta \leq \hat{\beta}$, all workers choose to be high-skilled if there were no discrimination. To analyze how does discrimination have impact on workers' expected payoff from search hence the skill investment incentives, it is important to first study the wage posting subgame with discrimination given a skill distribution. Without loss of generality, we start with the case where both groups choose to be high skilled. In the next section, we review the results from LMD (2005), where they study the case with discrimination but no difference in workers' skill levels (or productivity).

3.3.2 Existing results revisited and reinterpreted

In a context where there are two groups of workers with identical productivity (skill level) and firms strongly prefer group a to group b. LMD (2005) show that any subgame-perfect competitive equilibrium (SPCE) is separating.

Separating equilibrium. LMD (2005) show that there is no wage to which both groups of job seekers apply. More precisely, no wage can maximize firms' profit while attracting both groups of workers simultaneously (with the expected payoff from application attaining the corresponding market payoff). The equilibrium is separating. That is, there are some firms posting a higher level of wage attracting only the preferred group a, whereas the rest of firms offering a lower wage which is applied only by the discriminated group b (see Proposition 2 in LMD (2005)). Notice that the discriminated group have always the choice of applying to the high wage firms, however, they choose not to do so at all, because they anticipate discrimination in these firms. The most essential results of LMD (2005) are summarized as follows:

(i) At the firms attracting group a workers, expected profit for the firms and the expected payoff for the workers in the equilibrium are

$$\pi^S_{aH} = \left(1 - e^{-q^S_{aH}} - q^S_{aH} e^{-q^S_{aH}}\right) \times y_H$$

$$U^S_{aH} = e^{-q^S_{aH}} \times y_H$$

$$w^S_{aH} = \frac{q^S_{aH} e^{-q^S_{aH}}}{1-e^{-q^S_{aH}}} \times y_H$$

ii) At the firms attracting group b workers, expected profit for the firms and the expected payoff for the workers in the equilibrium are

$$\pi^S_{bH} = \left(1 - e^{-q^S_{bH}}\right)\left(1 - e^{-q^S_{aH}}\right) \times y_H,$$

$$w^S_{bH} = U^S_{aH}$$

$$U^S_{bH} = \frac{1-e^{-q^S_{bH}}}{q^S_{bH}} w^S_{bH}$$

Separating equilibrium requires that firms be indifferent between attracting group a and group b workers, that is

$$\pi^S_{aH} = \pi^S_{bH}$$

which helps to determine q^S_{bH} and q^S_{aH} jointly.

(iii) Furthermore, we have $w^S_{aH} > w^S_{bH}$. $q^S_{aH} > \beta > q^S_{bH}$, both q^S_{aH} and q^S_{bH} are increasing in β and γ, and both q^S_{aH} and q^S_{bH} are independent of y_H.

We make some important remarks on the features of the separating equilibrium. Firstly, the resulted equilibrium allocations are incentive compatible. For any particular *bH* job seeker, by deviating to applying for w^S_{aH}, the best they can get is $e^{-q^S_{aH}} \times w^S_{aH}$ (when none of the group *a* shows up in the firm which this *bH* worker deviates to apply to). However, this deviating payoff is strictly lower than sticking to applying to w^S_{bH} owing to the following relationship:

$$\begin{aligned} e^{-q^S_{aH}} \times w^S_{aH} &= e^{-q^S_{aH}} \times \frac{q^S_a \times e^{-q^S_a}}{1-e^{-q^S_a}} \times y_H = \frac{q^S_a \times e^{-q^S_a}}{1-e^{-q^S_a}} \times U^S_{aH} \\ &= e^{-q^S_{bH}} U^S_{aH} = e^{-q^S_{bH}} w^S_{bH} < \frac{1-e^{-q^S_{bH}}}{q^S_{bH}} w^S_{bH} = U^S_{bH} \end{aligned}$$

. As for any particular aH job seeker, by deviating to w^S_{bH}, the best they can get is $w^S_{bH} = U^S_{aH}$, which is as good as what he could get if he does not deviate. Secondly, we do not have the reservation wage structure. A reservation wage structure requires that workers apply to any wage which gives them an expected payoff higher than certain reservation value, in our case it would imply group bH should apply to both the low and high wages, however, it is not the case. This is because the expected payoff from applying to the high wage is a strictly dominated strategy for group b: the expected payoff from applying to high wages is too low to match their expected market payoff U^S_{bH}. Following are several noteworthy properties of such an equilibrium.

Results from LMD (2005)*: Compared to the context without discrimination, (1) Both groups have lower expected payoff. (2) All firms earn higher profits. (3) The expected payoff of group a and group b are such that* $U^S_{aH} > U^S_{bH}$.

Group b workers are worse off, because of firms' discriminatory hiring norm. Anticipating discrimination, group b demand lower expected payoff, which makes them cheaper to be employed. This in turn increases firms' market power in hiring group a. Group a understand that if they demand high wages, firms will threat to switch to hiring group b workers instead. Hence group a workers demand also low wages, and are worse off too. Now more about firms. Apart from the mechanisms just described, firms are able to earn high profits because in the regime with discrimination market is segmented, which allows the firms to face less competition in each segment. As a general remark, discrimination enables firms to extract higher profit by holding up job seekers' skill investment and providing all the job seekers lower expected payoff. Furthermore, since $U^S_{aH} > U^S_{bH}$, it suggests that the group b job seekers, being discriminated, are hurt to a larger extent. So that group b' incentive of skill investment is distorted further downwards. We enter more detailed discussions in the following section.

3.3.3 Analysis under our context

In the last section, we interpreted the equilibrium of the wage posting subgame given that all workers choose to be high skilled. In this section, we study how discrimination leads to different incentives of skill investment for these two groups respectively, and attempt to find the corresponding equilibrium.

An important observation is that the skill decision for group a and group b is strategic, and this is a direct consequence of the coexistence of ranking through the productivity-dependent (skill) and -independent traits (discrimination). Ranking by skills requires that the high skilled workers have the priority; While ranking by productivity-independent traits means that group a have the priority. Although multidimensional characteristics are involved, these two ranking schedules yield a unique market hierarchy:

$$aH \succ bH \succ aL \succ bL$$

. It reads as follows: high skilled group a (aH) are preferred to high skilled group b (bH), who are preferred to low skilled group a (aL), who are then preferred to low skilled group b (bL). How does the strategic interdependence arise? Take group a as an example for explanation. Although they are always ranked prior to group b due to discrimination, whenever they contemplate to lower skill investment, they understand that they will be ranked behind the high skilled group b; then the term $e^{-q_{bH}}$ which captures the competition from bH will appear in their payoffs.

Given the strategic interdependence in payoffs from skill investment, we adopt Nash equilibrium as the solution concept for the skill investment game. We proceed under Notion 1, and regard a whole group as making decision collectively and simultaneously, then all the workers in one group either end up high or low skilled. And we will focus on pure strategy of each group. Let α_g be the probability of group g's choosing to be high skilled, for $g = a$ or b. We have either $\alpha_g = 1$ or $\alpha_g = 0$. We have the following four possibilities as equilibrium:

- (P1) $\alpha_a = 1$ and $\alpha_b = 1$: group a - high, group b - high
- (P2) $\alpha_a = 1$ and $\alpha_b = 0$: group a - high, group b - low
- (P3) $\alpha_a = 0$ and $\alpha_b = 1$: group a - low, group b - high
- (P4) $\alpha_a = 0$ and $\alpha_b = 0$: group a - low, group b - low

To decide the skill investment, workers take into account firms' best response in the wage posting stage to infer the expected payoff from application, and compare the payoffs net of the cost of skill investment. In the wage posting subgame, when facing all workers with identical skill level (as in case (P1) and (P4)), firms' optimal strategy is the same as stated in LMD (2005); that is, some firms post a higher wage which is only applied by group a, whereas the rest post a lower wage which is only applied by group b. When there are both low and high skilled workers (as in case (P2) and (P3)), firms post wages conditional on skill level, and it is optimal for firms to attract both skill levels and rank the high skilled in priority to low skilled, as in Shi (2006). We then proceed to find workers' best response in the skill investment stage, and in turn the equilibrium in this discriminatory context with skill investment. We will use P1, P2, P3, P4 as the superscript for corresponding equilibrium allocations. When $\alpha_a = 1$ and $\alpha_b = 1$, workers are composed of type aH and bH. Firms post wages separately. We have the following payoffs for aH and bH respectively,

$$U_{aH}^{P1} = e^{-q_{aH}^{P1}} y_H - E_H$$

and

$$U_{bH}^{P1} = \frac{1-e^{-q_{bH}^{P1}}}{q_{bH}^{P1}} e^{-q_{aH}^{P1}} y_H - E_H$$

where we recall that $q_{aH}^S = q_{bH}^{P1} < \beta < q_{aH}^{P1} = q_{bH}^S$.

When $\alpha_a = 1$ and $\alpha_b = 0$, workers are composed of type aH and bL. Firms post wages conditional on skills, attract both types at the same time, and rank the high skilled in priority to the low skilled. As for the payoffs, we have

$$U_{aH}^{P2} = e^{-q_{aH}^{P2}} (y_H - y_L) + e^{-q_{aH}^{P2} - q_{bL}^{P2}} y_L - E_H$$

and

$$U_{bL}^{P2} = e^{-q_{aH}^{P2} - q_{bL}^{P2}} y_L - E_L$$

where $q_{aH}^{P2} = \gamma\beta$, $q_{bL}^{P2} = (1-\gamma)\beta$, and $q_{aH}^{P2} + q_{bL}^{P2} = \beta$.

When $\alpha_a = 0$ and $\alpha_b = 1$, workers are composed of different skill levels, aL and bH, we have similarly

$$U_{aL}^{P3} = e^{-q_{aL}^{P3} - q_{bH}^{P3}} y_L - E_L$$

and

$$U_{bH}^{P3} = e^{-q_{bH}^{P3}} (y_H - y_L) + e^{-q_{aL}^{P3} - q_{bH}^{P3}} y_L - E_H$$

where $q_{aL}^{P3} = \gamma\beta$, $q_{bH}^{P3} = (1-\gamma)\beta$, and $q_{aL}^{P3} + q_{bH}^{P3} = \beta$.

When $\alpha_a = 0$ and $\alpha_b = 0$, workers are composed of type aL and bL. Both are of same skill level, firms will discriminate and post wages as in LMD (2005), and workers' payoffs are

$$U_{aL}^{P4} = e^{-q_{aL}^{P4}} y_L - E_H$$

and

$$U_{bL}^{P4} = \frac{1-e^{-q_{bL}^{P4}}}{q_{bL}^{P4}} e^{-q_{aL}^{P4}} y_L - E_L$$

Since the solution of q in LMD (2005) is independent of y_H and y_L, we have $q_{bH}^{P1} = q_{bL}^{P4} < \beta < q_{aL}^{P4} = q_{bH}^{P1}$. The payoff matrix is as follows

	bH	bL
aH	bH: $\frac{1-e^{-q_{bH}^{P1}}}{q_{bH}^{P1}} e^{-q_{aH}^{P1}} y_H - E_H$ aH: $e^{-q_{aH}^{P1}} y_H - E_H$	bL: $e^{-q_{aH}^{P2}-q_{bL}^{P2}} y_L - E_L$ aH: $e^{-q_{aH}^{P2}} (y_H - y_L) + e^{-q_{aH}^{P2}-q_{bL}^{P2}} y_L - E_H$
aL	bH: $e^{-q_{bH}^{P3}} (y_H - y_L) + e^{-q_{aL}^{P3}-q_{bH}^{P3}} y_L - E_H$ aL: $e^{-q_{aL}^{P3}-q_{bH}^{P3}} y_L - E_L$	bL: $\frac{1-e^{-q_{bL}^{P4}}}{q_{bL}^{P4}} e^{-q_{aL}^{P4}} y_L - E_L$ aL: $e^{-q_{aL}^{P4}} y_L - E_H$

A pure-strategy Nash equilibrium on skill investment consists of a profile of actions with the property that no single group as a whole can achieve a higher payoff by unilateral deviation. The equilibrium depends on the value of β. For example, holding group a high skilled, when we decide whether group b choose to be high or low skill, we must compare U_{bH}^{P1} and U_{bL}^{P2}. And we find that there exists at least one threshold $\hat{\beta}_2$ which is determined by $U_{bH}^{P1} = U_{bL}^{P2}$ and is such that $\hat{\beta}_2 < \hat{\beta}$; this implies that depending on different values of β, the group b may choose high or low skill, leading to potentially different equilibrium. For tractability, we should introduce the following two conditions which guarantee that any such threshold as $\hat{\beta}_2$ is unique:[9]

Assumption 3.1 The equality $e^{-q_{aH}^{P1}(\hat{\beta}_1)} y_H - E_H = e^{-\hat{\beta}_1} y_L - E_L$ *admits a unique solution* $\hat{\beta}_1$.

Assumption 3.2 The equality $\frac{1-e^{-q_{bH}^{P1}(\hat{\beta}_2)}}{q_{bH}^{P1}(\hat{\beta}_2)} e^{-q_{aH}^{P1}(\hat{\beta}_2)} y_H - E_H = e^{-\hat{\beta}_2} y_L - E_L$ *admits a unique solution* $\hat{\beta}_2$.

The results on the equilibrium are summarized as follows:

Proposition 3. *Under the above assumptions, there exist two thresholds* $\hat{\beta}_2$ *and* $\hat{\beta}_1$ *with* $0 < \hat{\beta}_2 < \hat{\beta}_1 < \hat{\beta}$, *such that*

(1) When $0 < \beta < \hat{\beta}_2$, *there exists a unique equilibrium in which both group a and group b invest in high skill,* (aH, bH). *At* $\beta = \hat{\beta}_2$, *both* (aH, bH) *and* (aH, bL) *can be equilibrium.*

(2) When $\hat{\beta}_2 < \beta < \hat{\beta}_1$, *there exists a unique equilibrium in which group a invest in high skill while group b in low skill,* (aH, bL). *At* $\beta = \hat{\beta}_1$, *both* (aL, bH) *and* (aH, bL) *can be equilibrium.*

(3) When $\hat{\beta}_1 < \beta \leq \hat{\beta}$, *there exist multiple equilibria. Either group a invest in high skill and group b in low skill, or group a invest in low skill and group b invest in high skill,* (aL, bH) *or* (aH, bL).

Proof. In the Appendix.

[9]When y_H is sufficiently large compared to y_L, these two conditions are satisfied. Take $\hat{\beta}_1$ for example. The related assumption is satisfied when $\frac{e^{-q_{aH}^{P1}(\beta)} \frac{dq_{aH}^{P1}}{d\beta}}{e^{-\beta}} < \frac{y_H}{y_L}$. Since q_{aH}^{P1} does not depend on y_L and y_H, there exists always a pair of y_L and y_H such that this condition is satisfied.

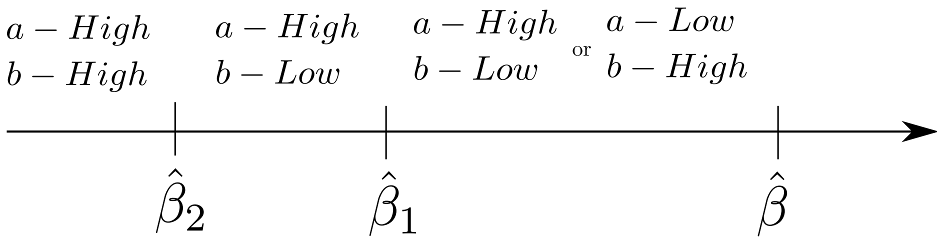

Fig. 3.1 Equilibrium depending on β

In fact, if one group choose to be low skilled, the best response of the other group is always to be high skilled, while the best response to the other's high skill choice depends on the two thresholds. Furthermore, the rise of market tightness β makes workers have stronger incentive to deviate from high skill, and group *b* is more prone to deviate compared to group *a*, in the sense that the threshold of β at which group *b* begins to contemplate to invest in low skill is lower compared to group *a*. Interestingly, for values of β close to $\hat{\beta}$, there exists an equilibrium where the preferred group *a* choose low skill, while the discriminated group *b* choose high skill. The results are summarized in Figure 3.1. We have the following results on the comparison of workers' expected payoff and firms' profits compared to the case without discrimination.

Corollary. *Compared to the case without discrimination,*

(1) In (aH, bH) equilibrium, firms always earn higher expected profits; In (aH, bL) equilibrium and (aL, bH), firms earn lower expected profits.

(2) In (aH, bH) equilibrium, both aH and bH workers earn lower expected payoff; In (aH, bL) equilibrium, group aH (group bL) earn higher (lower) expected payoff; in (aL, bH) equilibrium, group bH (group aL) earn higher (lower) expected payoff.

Proof. In the Appendix.

This corollary tells that firms can be worse off with discrimination. Indeed, when workers anticipate discrimination, their investment incentive may be downwards distorted, and some group may end up choosing to underinvest. In equilibrium, whenever one group underinvest and the other group remain high skilled, the firms turn out to earn lower expected profits compared to the case without discrimination. This is simply due to the fact that underinvestment in skills discretely drags down the average productivity of the economy.

3.4 Comparison with fixed sharing rule (wage bargaining)

In this section, we shut down the channel through which firms use wages to influence workers' choices on applications, and examine whether the inefficiency can be alleviated. Notice that in this section, the workers only choose the amount of skills to obtain, not where to search.

Consider an economy with the same discriminatory ranking as previous, but the wage is determined by ex post bargaining after a job seeker meets an employer. The timing of the economy now is as follows: firstly, workers decide skill levels simultaneously; secondly, workers and firms get matched according to the matching technology; thirdly, the matched worker and firm bargain a la Nash to determine how to share the output y. The simplest form of Nash bargaining widely used in literature is equivalent to a fixed sharing rule of output. If we denote the bargaining power for all workers as ψ, then from the output y_t, workers receive ψy_t, and firms receive $(1-\psi)y_t$. We focus on the case where ψ is the same for both skill levels, otherwise there is too much degree of freedom.

For the ease of comparison, we require that the matching technology here is the same as in previous section. The hiring norm is as previous Group $aH \succ$ Group $bH \succ$ Group $aL \succ$ Group bL. The corresponding employment probability for different types of workers is inherited, so that the employment probability is respectively $\frac{1-e^{-q_{aH}}}{q_{aH}}$ for aH, $e^{-q_{aH}}\frac{1-e^{-q_{bH}}}{q_{bH}}$ for bH, $e^{-q_{aH}-q_{bH}}\frac{1-e^{-q_{aL}}}{q_{aL}}$ for aL, and $e^{-q_{aH}-q_{bH}-q_{aL}}\frac{1-e^{-q_{bL}}}{q_{bL}}$ for bH. The expected payoff (after skill investment) is just the employment probability times ψy_t; for example, we have for aL people $e^{-q_{aH}-q_{bH}}\frac{1-e^{-q_{aL}}}{q_{aL}} \times \psi y_L$ as the expected payoff from search.

Now, we specify the expected queue lengths q_{aH}, q_{bH}, q_{aL} and q_{bL} parametrically. Recall the definition of queue length is $\frac{\text{nb. of workers}}{\text{nb. of vacancies}}$, then we have $q_{aH} = \gamma\eta_a\beta$ for aH workers, $q_{bH} = (1-\gamma)\eta_b\beta$ for bH workers, $q_{aL} = \gamma(1-\eta_a)\beta$ for aL workers, and $q_{bL} = (1-\gamma)(1-\eta_b)\beta$ for bL workers, where η_a represents the fraction of high skilled group a, and η_b the fraction of high skilled group b. The values of η_a and η_b depend on the comparison between the expected payoff from investing in high or low skill:

$$\eta_a \begin{cases} = 1 & \text{if } \frac{1-e^{-q_{aH}}}{q_{aH}} \times \psi y_H - E_H > e^{-q_{aH}-q_{bH}}\frac{1-e^{-q_{aL}}}{q_{aL}} \times \psi y_L - E_L \\ = 0 & \text{if } \frac{1-e^{-q_{aH}}}{q_{aH}} \times \psi y_H - E_H < e^{-q_{aH}-q_{bH}}\frac{1-e^{-q_{aL}}}{q_{aL}} \times \psi y_L - E_L \\ \in \{0,1\} & \text{if } \frac{1-e^{-q_{aH}}}{q_{aH}} \times \psi y_H - E_H = e^{-q_{aH}-q_{bH}}\frac{1-e^{-q_{aL}}}{q_{aL}} \times \psi y_L - E_L \end{cases}$$

and

$$\eta_b \begin{cases} = 1 & \text{if } e^{-q_{aH}} \frac{1-e^{-q_{bH}}}{q_{bH}} \times \psi y_H - E_H > e^{-q_{aH}-q_{bH}-q_{aL}} \frac{1-e^{-q_{bL}}}{q_{bL}} \times \psi y_L - E_L \\ = 0 & \text{if } e^{-q_{aH}} \frac{1-e^{-q_{bH}}}{q_{bH}} \times \psi y_H - E_H < e^{-q_{aH}-q_{bH}-q_{aL}} \frac{1-e^{-q_{bL}}}{q_{bL}} \times \psi y_L - E_L \\ \in \{0,1\} & \text{if } e^{-q_{aH}} \frac{1-e^{-q_{bH}}}{q_{bH}} \times \psi y_H - E_H = e^{-q_{aH}-q_{bH}-q_{aL}} \frac{1-e^{-q_{bL}}}{q_{bL}} \times \psi y_L - E_L \end{cases}$$

To keep consistency with the previous section (under Notion 1), we require that whenever indifferent, the whole group will choose either high or low skill, so that η is either 0 or 1 in that case. We will also assume that the group a are the majority: $\gamma \geq \frac{1}{2}$.[10]

We focus on Nash equilibrium as the solution concept. In the current context, only workers make skill investment decisions, firms do not post wages since ψ is exogenous. We consider each group, i.e. group a or group b, as a whole when they are making decisions. Each group of workers invest in skills simultaneously. Due to the discriminatory rule, the payoffs from skill investment for different groups of workers are interdependent. This renders the skill investment strategic. The following proposition helps explain how workers' expected payoffs vary with respect to ψ:

Proposition 4. Let $\gamma \geq \frac{1}{2}$ so that group a is the majority. There are four thresholds $\hat{\psi}_{aL,b} \leq \hat{\psi}_{bH,a} < \hat{\psi}_{bL,a} < \hat{\psi}_{aH,b}$, such that

(1) For $\psi \in \left[0, \hat{\psi}_{aL,b}\right)$, the unique pure strategy nash equilibrium is (aL, bL); for $\psi = \hat{\psi}_{aL,b}$, the equilibrium can be (aL, bL) or (aL, bH). (2) For $\psi \in \left(\hat{\psi}_{aL,b}, \hat{\psi}_{bH,a}\right)$, the unique pure strategy nash equilibrium is (aL, bH); (3) For $\psi \in \left(\hat{\psi}_{bH,a}, \hat{\psi}_{bL,a}\right)$, there is no pure strategy nash equilibrium; for $\psi = \hat{\psi}_{bL,a}$, the equilibrium is (aH, bL); (4) For $\psi \in \left(\hat{\psi}_{bL,a}, \hat{\psi}_{aH,b}\right)$, the unique pure strategy nash equilibrium is (aH, bL); at the point $\psi = \hat{\psi}_{aH,b}$, the equilibrium can be (aH, bL) or (aH, bH). (5) For $\psi \in \left(\hat{\psi}_{aH,b}, 1\right)$, the unique pure strategy nash equilibrium is (aH, bH). (6) Define the threshold $\hat{\psi}$ of skill investment without discrimination as $\hat{\psi} y_H \frac{1-e^{-\beta}}{\beta} - E_H = \hat{\psi} y_L \frac{1-e^{-\beta}}{\beta} - E_L$; then $\hat{\psi}_{bL,a} < \hat{\psi} < \hat{\psi}_{aH,b}$.

Proof. In the Appendix.

It is not always true that the group a workers are always better off. Notably, there is an equilibrium similar as before where group a underinvest: It could be observed from the payoff matrix (provided in the proof of Proposition 4) that in the region $\psi \in \left(\hat{\psi}_{aL,b}, \hat{\psi}_{bH,a}\right)$ group a (group b) workers obtain lower (higher) expected payoff compared to the case without

[10] We think this case with $\gamma \geq \frac{1}{2}$ is more empirically relevant, when we are talking about gender and racial discrimination for example. The case $\gamma < \frac{1}{2}$ could be also analogously derived according to the proof of the following proposition.

discrimination. This is also closely related to the fact that when $\gamma \geq \frac{1}{2}$ the within-group competition in group a is fiercer. In general, firms' payoff is written as follows:

$$\pi = \begin{cases} \left(1 - e^{-q_{aH} - q_{bH}}\right)(1-\psi)\, y_H & \text{if all high skilled} \\ \begin{array}{c} \left(1 - e^{-q_{aH} - q_{bH}}\right)(1-\psi)\, y_H \\ + e^{-q_{aH} - q_{bH}}\left(1 - e^{-q_{aL} - q_{bL}}\right)(1-\psi)\, y_L \end{array} & \text{if both high and low skilled} \\ \left(1 - e^{-q_{aL} - q_{bL}}\right)(1-\psi)\, y_L & \text{if all low skilled} \end{cases}$$

According to Proposition 4, we can determine the exact values of the queue lengths in the expression. Firms' profit will be piecewise monotone because although ψ increases continuously, the skill composition hence the average productivity of the market improves discretely with respect to this bargaining power. The fact that $\hat{\psi}_{bL,a} < \hat{\psi} < \hat{\psi}_{aH,b}$ suggests that although firms can gather higher profits for $\psi < \hat{\psi}$, they encounter loss for $\psi \geq \hat{\psi}$ compared to the case without discrimination. The reason is that strategic competition between the group a and group b deters the discriminated group's skill investment decision (in the sense that group b may still choose to be low skilled when ψ is sufficiently high), which pulls down the market's average productivity and makes firms' expected profit dim.

It is interesting to notice that our simple result that discrimination is costly for firms at high skilled sector (when wages are bargained) questions the plausibility of key assumption of Merlino (2012) that "there is more discrimination in the high technology sector". Although Merlino (2012) mentioned bunches of empirical evidence in support of this assumption[11], our simple result suggest that firms are simply better off not discriminating when wages are principally bargained, since the loss in profit from discriminating in the high skilled sector may surpass the gain from discriminating in the low skilled sector. All in all, the key difference between wage posting and wage bargaining is that the ex post wage now exogenously pegs on the productivity, and firms can no longer manipulate their market power by translating their discriminatory preference into constantly lower wages.

3.5 Discussion

Free entry. LMD (2005) have shown that their economy under discrimination with workers' identical in productivity can be generalized to take into account firms' free entry. Specifically, we consider a stage where firms sink capital after observing workers' skills. Each firm has different capital cost with $C_1 < C_2 < ... < C_M < y_L$. Then firms which earn expected

[11] See Merlino (2012) page 4 for more relevant reference.

non-positive profits after the reduction of capital cost would simply not enter into the market. In the paper, we observe that the equilibrium is unique with respect to β, which has a one-one relationship with M - the number of firms in the market, so that the results in the paper could carry through with free entry. All firms in the market expect positive net profits. When there are different skill groups, this result could also carry through, because the equilibrium profit of firms is still an increasing function of β.

Heterogeneity in skill investment cost. Some preliminary attempts from us suggest that our context could be generalized to a situation where workers are heterogenous in their skill investment cost (although more complicated): let the low skill investment cost be zero ($E_L = 0$) for all workers, and the high skill investment cost be, for simplicity, of two values $E_{H,1} < E_{H,2}$; there are still two levels of productivity: y_L and y_H. Focus on the corresponding $\hat{\beta}$ and define it as $\hat{\beta} = \log \frac{y_H - y_L}{E_{H,2} - E_L}$. If the contracts can be contingent on $E_{H,1}$ and $E_{H,2}$, the submarkets for type $E_{H,1}$ workers and type $E_{H,2}$ workers are separated, and all the results in the paper carry through for the workers of cost $E_{H,2}$; as for the workers of cost $E_{H,1}$, their skill investment cost is lower, hence they have stronger incentive to remain high skilled; then for values of β close to $\hat{\beta} = \log \frac{y_H - y_L}{E_{H,2} - E_L}$, some equilibrium which exists in the $E_{H,2}$ submarket may not exist in the $E_{H,1}$ submarket. If the contracts can not be contingent on $E_{H,1}$ and $E_{H,2}$, both type $E_{H,1}$ and type $E_{H,2}$ are in the same market and will compete; as a result, there may exist a region of β where both high skilled group a and group b, as well as both low skilled group a and group b, exist at the same time. The extent of the skill investment game is in turn larger, because, for example, a particular group "$a, E_{H,1}$"'s skill investment decision should be a best response of other groups: "$a, E_{H,2}$", "$b, E_{H,1}$", and "$b, E_{H,2}$". If heterogeneity in skill investment cost is managed, it is possible to extend the model to multiple skill levels. Shi (2006) shows that in such a model with multiple skill levels free of discrimination, the result that firms always rank the high skilled workers in priority to the workers with lower skills can be generalized. The difficulty under the context with discrimination, as just stated, is on the extent of the game.

3.6 Conclusion

In this paper, we study a holdup problem where firms can use discriminatory hiring norms to extract higher than socially optimal profits. We find that when firms rank workers according to both productivity-dependent and productivity-independent characteristics, skill investment becomes strategic between the discriminated and the favored group. In case wages are posted, we suggest that depending on the market tightness there may be equilibrium or multiple equilibria on skill investment; in some equilibrium the discriminated group can obtain higher expected payoff compared to the case without discrimination[12] and firms can be worse off. We also consider fixed sharing rule (bargained wage) and make a comparison. Similar equilibrium, where favored group underinvest while the discriminated group remain high skilled, exists; however, the discriminated group are in general worse off compared to the case without discrimination in the sense that they may still choose to underinvest when ψ is sufficiently high. Firms' profits are piecewise monotone because the skill composition hence the average productivity of the market improves discretely with respect to the bargaining power, and profit loss may be incurred with discrimination within an intermediate range of bargaining power.

3.7 Appendix

A1. Derivation of matching probabilities.

We now derive a job seeker's matching probability and expected payoff.

Job seekers. Having observed all the wage $\mathbf{w} = \left\{w^1, w^2, ..., w^M\right\}$ announced by the firms, job seekers choose which firm (or wage) to visit (or to apply for). Consider a particular job seeker i's problem, where $i \in \{1, 2, ..., N\}$. This job seeker thinks in the following way: Suppose I visit firm j, then conditional on the fact that my application is sent to j, what is the probability that I could be employed? It depends upon the number of *the other job seekers* who also send their job application to the same firm competing with me on this job in firm j. This number (of *the other job seekers*) is a random variable which has a realisation from the set $\{0, 1, ..., N-1\}$ and has a Binomial distribution. To see why it is the case, we use k to represent the realized number of competitors. If $k = 0$, which happens with probability $\left(1-\theta^j\right)^{N-1}$, then the job seeker i will be chosen by the firm with probability 1, because this job seeker is the only candidate. If $k = 1$, which happens with probability $(N-1) \times \left(\theta^j\right)^1 \left(1-\theta^j\right)^{(N-1)-1}$, this job seeker i will be chosen by the firm with probability $\frac{1}{2}$, because now the firm receives two applications, hence has two candidates, among whom i is one. Generalising, if $k = \hat{k}$, which happens with probability $C_{N-1}^{\hat{k}} \times \left(\theta^j\right)^{\hat{k}} \left(1-\theta^j\right)^{N-1-\hat{k}}$, then this job seeker i will be chosen with probability $\frac{1}{\hat{k}+1}$, because the firm j has $\hat{k}+1$ candidates at disposal.

[12] Recall that without discrimination, it is socially optimal.

The employment probability for the workers is $\sum_{k=0}^{N-1} C_{N-1}^{k} \left(\theta^{j}\right)^{k} \left(1-\theta^{j}\right)^{N-1-k} \frac{1}{k+1}$. This expression could be simplified to $\frac{1-\left(1-\theta^{j}\right)^{N}}{N\theta^{j}}$ [13]. Hence the job seeker's expected pay off is $\frac{1-\left(1-\theta^{j}\right)^{N}}{N\theta^{j}} \times w^{j}$.

A2. Proofs of propositions

Proposition 1. (return to skills) *Given the return to skill ratio $\frac{y_H-y_L}{E_H-E_L}$, define $\hat{\beta}$ as $\frac{y_H-y_L}{E_H-E_L} = e^{\hat{\beta}}$.*

(i) when $0 < \beta \leq \hat{\beta}$, the unique equilibrium is such that all job seekers choose high skill, i.e. $\alpha^* = 1$.

(ii) when $\beta > \hat{\beta}$, the equilibrium consists of a unique value $\alpha^* \in (0,1)$ which satisfies $\frac{y_H-y_L}{E_H-E_L} = e^{\alpha^*\beta}$.

(iii) when $\frac{y_H-y_L}{E_H-E_L} \leq 1$ such that $\frac{y_H-y_L}{E_H-E_L} = e^{\hat{\beta}}$ has no positive solution on $\hat{\beta}$, the unique equilibrium is $\alpha^* = 0$.

Proof. We will prove only case (i) while the proof of case (ii) and (iii) are highly similar. Notice first that $\frac{y_H-y_L}{E_H-E_L} \geq e^{\beta}$ is equivalent to $e^{-\beta}y_H - E_H \geq e^{-\beta}y_L - E_L$.

We prove firstly that the deviation to low skill is not optimal. By this, we prove that a proportion ε of workers' deviating to low skill is suboptimal. And it suffices to show that after deviation, the deviator can not get higher expected payoff. Before deviation, the expected payoff is $e^{-q_H^*}y_H - E_H$, where $q_H^* = \beta$. After deviation, the expected payoff becomes $e^{-q_H^D-q_L^D}y_L - E_L$, where $q_H^D + q_L^D = \beta$. However, under the condition $e^{-\beta}y_H - E_H \geq e^{-\beta}y_L - E_L$, the expected payoff after deviation is weakly lower.

For the uniqueness. We should furthermore show that for the case of $\alpha = 0$ and $\alpha \in (0,1)$, there will be profitable deviation. When $\alpha = 0$, the expected payoff from search is $e^{-\beta}y_L - E_L$. If there is a fraction ε deviating to high skilled, then the expected income for the deviator becomes $e^{-\varepsilon\beta}\left(y_H - y_L\right) + e^{-\beta}y_L - E_H$. Then this expected payoff after deviation is greater than the expected payoff before deviation because. $e^{-\varepsilon\beta}\left(y_H - y_L\right) > e^{-\beta}\left(y_H - y_L\right) \geq E_H - E_L$. So the deviation is profitable for the deviators. When $\alpha \in (0,1)$, the expected income from search is $e^{-\beta}y_L - E_L$ for the low skilled, and $e^{-\hat{\alpha}\beta}\left(y_H - y_L\right) + e^{-\beta}y_L - E_H$ for the type H job seekers, where $\hat{\alpha}$ should be pinned down by workers' indifference condition $e^{-\hat{\alpha}\beta}\left(y_H - y_L\right) = E_H - E_L$. However, this condition is incompatible for any $\alpha < 1$ with our condition $e^{-\beta}y_H - E_H \geq e^{-\beta}y_L - E_L$. So that it is impossible that job seekers are indifferent from being high or low skilled.

[13] One way of deriving it could be seen in Melanie Cao & Shouyong Shi, 2000. "Coordination, matching, and wages". It could also be checked by change of variable, which is also represented in the Appendix.

All in all, we have proved that when the configuration of parameters is such that $e^{-\beta}y_H - E_H > e^{-\beta}y_L - E_L$, the only equilibrium is all the job seekers choose to obtain high skill, i.e. $\alpha^* = 1$. Q.E.D.

Proposition 2. *The equilibrium labor allocation and skill investment choice are socially optimal.*

Proof. We derive the proof for Notion 2, the proof for Notion 1 can be analogously derived. By definition of q, we have $q_H = \frac{N_H}{M}$ and $q_L = \frac{N_L}{M}$ - all firms will attract both skill types. Since $N_H = \alpha N$ and $N_L = (1-\alpha)N$, we have $q_H = \alpha\beta$ and $q_L = (1-\alpha)\beta$. When $\alpha = 1$ or $\alpha = 0$, there is only one skill level present in the market; when $\alpha \in (0,1)$, there are both high and low skilled. Define a priority rule $R \in [0,1]$, which is the probability of choosing high skilled job applicants when both high and low skilled are present in the same firm. The planner chooses α, which hence determines q_H and q_L, to maximize the aggregate output

$$
\begin{gathered}
M \times \left[\begin{array}{c} (1-e^{-q_H})(R \times (1-e^{-q_L}) + e^{-q_L})y_H \\ +(1-e^{-q_L})((1-R)(1-e^{-q_H}) + e^{-q_H})y_L \end{array} \right] \\
-N \times (\alpha E_H + (1-\alpha)E_L)
\end{gathered}
$$

The objective program can be rearranged to the following way

$$
\begin{gathered}
M \times \left[\begin{array}{c} (1-e^{-q_L})(1-e^{-q_H}) \times [Ry_H + (1-R)y_L] \\ +e^{-q_L}y_H + e^{-q_H}y_L - e^{-\beta}y_H - e^{-\beta}y_L \end{array} \right] \\
-N \times (\alpha E_H + (1-\alpha)E_L)
\end{gathered}
$$

Whenever we are with corner solutions on α, i.e. $q_H = 0$ or $q_L = 0$, the expression does not depend on R. When the solution on α is interior, q_H and q_L are both positive. And if we maximize with respect to R, we have $(1-e^{-q_H})(1-e^{-q_L})(y_H - y_L) > 0$, so that setting $R^p = 1$ is the optimal choice. By doing so, we could further reduce the objective to

$$
\begin{gathered}
M \times [(1-e^{-q_H})y_H + e^{-q_H}(1-e^{-q_L})y_L] \\
-N \times (\alpha E_H + (1-\alpha)E_L)
\end{gathered}
$$

Now the derivative with respect to α yields

$$
\begin{gathered}
\overbrace{M\beta}^{=N} \times \left[e^{-\alpha\beta}(y_H - y_L)\right] \\
-N \times (E_H - E_L)
\end{gathered}
$$

where the first line represents the marginal gain from assigning 1% more workers to the high skilled section, and the second line represents the corresponding marginal loss. For values of β such that the marginal gain surpasses the marginal loss, the planner will set $\alpha = 1$, in which case a threshold $\hat{\beta}^p$ is determined by $e^{\hat{\beta}^p}(y_H - y_L) = E_H - E_L$, such that for values of β not larger than this threshold pinned down by the log-return to skills, the planner finds socially optimal to assign all workers to the high skilled sector. When the skill to return is such that $(y_H - y_L) < E_H - E_L$, for all values of β the marginal gain will be lower than the marginal loss, the planner will choose $\alpha^p = 0$. At last if β satisfies $\beta > \hat{\beta}^p$, such that the log-return to skill investment is not high enough to offset the market competition (captured by β), there is an $\alpha^p \in (0,1)$ (an interior solution) which is determined by $e^{-\alpha^p \beta}(y_H - y_L) = E_H - E_L$ such that the planner will recommend ex ante identical workers to randomize on skill choice and a fraction α^p will end up high skilled. It is straightforward to notice that the correspond values of $q_H^p = \alpha^p \beta$ and $q_L^p = (1 - \alpha^p)\beta$ correspond to the equilibrium allocation.

As a summary, we have shown that the threshold for skill investment $\hat{\beta}^p$ coincides with $\hat{\beta}$, and α^p conincides with α^*. Q.E.D.

Proposition 3. *Under the above assumptions, there exist two thresholds $\hat{\beta}_2$ and $\hat{\beta}_1$ with $0 < \hat{\beta}_2 < \hat{\beta}_1 < \hat{\beta}$, such that*

(1) When $0 < \beta < \hat{\beta}_2$, there exists a unique equilibrium in which both group a and group b invest in high skill, (aH, bH). At $\beta = \hat{\beta}_2$, both (aH, bH) and (aH, bL) can be equilibrium.

(2) When $\hat{\beta}_2 < \beta < \hat{\beta}_1$, there exists a unique equilibrium in which group a invest in high skill while group b in low skill, (aH, bL). At $\beta = \hat{\beta}_1$, both (aL, bH) and (aH, bL) can be equilibrium.

(3) When $\hat{\beta}_1 < \beta \leq \hat{\beta}$, there exist multiple equilibria. Either group a invest in high skill and group b in low skill, or group a invest in low skill and group b invest in high skill, (aL, bH) or (aH, bL).

Proof. Holding group b high skilled, group a's best response depends on the comparison between $U_{aH}^{P1}(\beta)$ and $U_{aL}^{P3}(\beta)$. On one hand, since $q_{aH}^{P1}(\beta) > \beta$, we have

$$U_{aH}^{P1}\left(\hat{\beta}\right) = e^{-q_{aH}^{P1}(\hat{\beta})} y_H - E_H < e^{-\hat{\beta}} y_H - E_H = e^{-\hat{\beta}} y_L - E_L = U_{aL}^{P3}\left(\hat{\beta}\right)$$

On the other hand, since q_{aH}^{P1} is increasing in β, when $\beta \to 0$, we have $U_{aH}^{P1}(\beta) \to y_H - E_H$ which is greater than $U_{aL}^{P3}(0) = y_L - E_L$. Due to the continuity of $U_{aH}^{P1}(\beta)$ and $U_{aL}^{P3}(\beta)$ on β, there exists at least one $\hat{\beta}_1 < \hat{\beta}$ such that $U_{aH}^{P1}\left(\hat{\beta}_1\right) = U_{aL}^{P3}\left(\hat{\beta}_1\right)$. According to the assumptions on the uniqueness of the intersection point, we have $U_{aH}^{P1}(\beta) > U_{aL}^{P3}(\beta)$ for $\beta < \hat{\beta}_1$, and $U_{aH}^{P1}(\beta) < U_{aL}^{P3}(\beta)$ for $\beta > \hat{\beta}_1$.

Holding group b low skilled, group a's best response depends on the comparison between U_{aH}^{P2} and U_{aL}^{P4}. It turns out that for all values of $\beta < \hat{\beta}$

$$U_{aH}^{P2} = e^{-\gamma\beta}(y_H - y_L) + e^{-\beta}y_L - E_H > e^{-\beta}y_H - E_H > e^{-q_{aH}^{P4}}y_L - E_L = U_{aL}^{P4}$$

where the last inequality is due to the fact that $e^{-\beta}y_H - E_H \geq e^{-\beta}y_L - E_L$ for all $\beta \leq \hat{\beta}$ and $q_{aH}^{P4} > \beta$, implying that choosing high skill is a dominant strategy for group a when group b choose low skill.

Holding group a high skilled, group b's best response depends on the comparison between U_{bH}^{P1} and U_{bL}^{P2}. On one hand, we have

$$U_{bH}^{P1}\left(\hat{\beta}\right) = \frac{1-e^{-q_{bH}^{P1}(\hat{\beta})}}{q_{bH}^{P1}(\hat{\beta})}e^{-q_{aH}^{P1}(\hat{\beta})}y_H - E_H < e^{-\hat{\beta}}y_H - E_H = e^{-\hat{\beta}}y_L - E_L = U_{bL}^{P2}\left(\hat{\beta}\right)$$

where the first inequality is due to the fact that $\frac{1-e^{-q_{bH}^{P1}(\hat{\beta})}}{q_{bH}^{P1}(\hat{\beta})}e^{-q_{aH}^{P1}(\hat{\beta})} < e^{-q_{aH}^{P1}(\hat{\beta})}$ and $q_{aH}^{P1}(\beta) > \beta$. On the other hand, since $q_{aH}^{P1}(\beta)$ and $q_{bH}^{P1}(\beta)$ are increasing in β, when $\beta \to 0$, we have $U_{bH}^{P1}(\beta) \to y_H - E_H$ which is greater than $U_{bL}^{P2}(0) = y_L - E_L$. Due to the continuity of $U_{bH}^{P1}(\beta)$ and $U_{bL}^{P2}(\beta)$ on β, there exists at least one $\hat{\beta}_2 < \hat{\beta}$ such that $U_{bH}^{P1}\left(\hat{\beta}_2\right) = U_{bL}^{P2}\left(\hat{\beta}_2\right)$. According to the assumptions on the uniqueness of the intersection point, we have $U_{bH}^{P1}\left(\hat{\beta}_2\right) > U_{bL}^{P2}\left(\hat{\beta}_2\right)$ for $\beta < \hat{\beta}_2$, and $U_{bH}^{P1}\left(\hat{\beta}_2\right) < U_{bL}^{P2}\left(\hat{\beta}_2\right)$ for $\beta > \hat{\beta}_2$.

Holding group a low skilled, group b's best response depends on the comparison between U_{bH}^{P3} and U_{bL}^{P4}. It turns out that for all values of $\beta \leq \hat{\beta}$ we have

$$\begin{aligned} U_{bH}^{P3} &= e^{-(1-\gamma)\beta}(y_H - y_L) + e^{-\beta}y_L - E_H \\ &> e^{-\beta}y_H - E_H \\ &\geq e^{-\beta}y_L - E_L \\ &> \frac{1-e^{-q_{bL}^{P4}(\beta)}}{q_{bL}^{P4}(\beta)}e^{-q_{aL}^{P4}(\beta)}y_L - E_L \end{aligned}$$

where the last inequality is due to the fact that $e^{-\beta} > e^{-q_{aL}^{P4}} > \frac{1-e^{-q_{bL}^{P4}(\beta)}}{q_{bL}^{P4}(\beta)}e^{-q_{aL}^{P4}(\beta)}$, implying that choosing high skill is a dominant strategy for group b when group a choose low skill.

To summarize, for values of $\beta \in \left(0, \hat{\beta}_2\right)$, both groups choosing high skill, i.e. (aH, bH), is the unique equilibrium; for $\beta \in \left(\hat{\beta}_2, \hat{\beta}_1\right)$, group a choosing high skill and group b choosing low skill, i.e. (aH, bL) is the unique equilibrium; When $\beta \in \left(\hat{\beta}_1, \hat{\beta}\right)$, both (aH, bL) and (aL, bH) are possible to appear as equilibrium. Q.E.D.

Corollary. *Compared to the case without discrimination,* *(1) In (aH, bH) equilibrium, firms always earn higher expected profits; in In (aH, bL) equilibrium and (aL, bH), firms earn lower expected profits.*

(2) In (aH, bH) equilibrium, both aH and bH workers earn lower expected payoff; in (aH, bL) equilibrium, group aH (group bL) earn higher (lower) expected payoff; in (aL, bH) equilibrium, group bH (group aL) earn higher (lower) expected payoff.

Proof. (1) In (aH, bH) equilibrium, the result is proved in LMD (2005). We prove the case for (aH, bL) equilibrium. In (aH, bL) equilibrium, firms' profits are

$$\pi^{P2}_{aH,bL} = \left(1 - e^{-q_H^{P2}} - q_H^{P2} e^{-q_H^{P2}}\right) y_H + \left[\left(q_H^{P2} + 1\right) e^{-q_H^{P2}} - e^{-\beta}(\beta + 1)\right] y_L$$

This term is smaller than the profit without discrimination, because

$$\begin{aligned}
\pi^{P2}_{aH,bL} &= \left(1 - e^{-q_H^{P2}} - q_H^{P2} e^{-q_H^{P2}}\right) y_H + \left[\left(q_H^{P2} + 1\right) e^{-q_H^{P2}} - e^{-\beta}(\beta + 1)\right] y_L \\
&< \left(1 - e^{-q_H^{P2}} - q_H^{P2} e^{-q_H^{P2}}\right) y_H + \left[\left(q_H^{P2} + 1\right) e^{-q_H^{P2}} - e^{-\beta}(\beta + 1)\right] y_H \\
&= \left(1 - e^{-\beta} - e^{-\beta}\beta\right) y_H
\end{aligned}$$

where the inequality uses the fact that $(x+1)e^{-x}$ is a decreasing function and $q_H^{P2} < \beta$. The proof for the case of (aL, bH) equilibrium can be analogously reproduced.

(2) For the case (aH, bH), it follows from LMD (2005). For the case of (aH, bL). We have

$$\begin{aligned}
U^{P2}_{aH} &= e^{-\gamma\beta}(y_H - y_L) + e^{-\beta} y_L - E_H \\
&> e^{-\beta} y_H - E_H \\
&> e^{-q^S_{aH}} y_H - E_H
\end{aligned}$$

where the first inequality comes from $e^{-\gamma\beta} - e^{-\beta} > 0$.

$$\text{and} \quad \begin{aligned}
U^{P2}_{bL} &= e^{-\beta} y_L - E_L \\
&< e^{-\beta} y_H - E_H
\end{aligned}$$

where the inequality comes from the fact that $\beta \leq \hat{\beta}$. The proof for the case of (aL, bH) equilibrium can be analogously reproduced. **Q.E.D.**

Proposition 4. *Let $\gamma \geq \frac{1}{2}$ so that group a is the majority. There are four thresholds $\hat{\psi}_{aL,b} \leq \hat{\psi}_{bH,a} < \hat{\psi}_{bL,a} < \hat{\psi}_{aH,b}$, such that*

(1) For $\psi \in [0, \hat{\psi}_{aL,b})$, the unique pure strategy nash equilibrium is (aL, bL); for $\psi = \hat{\psi}_{aL,b}$, the equilibrium can be (aL, bL) or (aL, bH). (2) For $\psi \in (\hat{\psi}_{aL,b}, \hat{\psi}_{bH,a})$, the unique pure strategy nash equilibrium is (aL, bH); (3) For $\psi \in (\hat{\psi}_{bH,a}, \hat{\psi}_{bL,a})$, there is no pure strategy nash equilibrium; for $\psi = \hat{\psi}_{bL,a}$, the equilibrium is (aH, bL); (4) For

$\psi \in \left(\hat{\psi}_{bL,a}, \hat{\psi}_{aH,b}\right)$, the unique pure strategy nash equilibrium is (aH, bL); at the point $\psi = \hat{\psi}_{aH,b}$, the equilibrium can be (aH, bL) or (aH, bH). (5) For $\psi \in \left(\hat{\psi}_{aH,b}, 1\right)$, the unique pure strategy nash equilibrium is (aH, bH). (6) Define the threshold $\hat{\psi}$ of skill investment without discrimination as $\hat{\psi} y_H \frac{1-e^{-\beta}}{\beta} - E_H = \hat{\psi} y_L \frac{1-e^{-\beta}}{\beta} - E_L$; then $\hat{\psi}_{bL,a} < \hat{\psi} < \hat{\psi}_{aH,b}$.

Proof. The payoff matrix is as follows.

	bH	bL
aH	bH: $e^{-\gamma\beta} \frac{1-e^{-(1-\gamma)\beta}}{(1-\gamma)\beta} \psi y_H - E_H$ aH: $\frac{1-e^{-\gamma\beta}}{\gamma\beta} \psi y_H - E_H$	bL: $e^{-\gamma\beta} \frac{1-e^{-(1-\gamma)\beta}}{(1-\gamma)\beta} \psi y_L - E_L$ aH: $\frac{1-e^{-\gamma\beta}}{\gamma\beta} \psi y_H - E_H$
aL	bH: $\frac{1-e^{-(1-\gamma)\beta}}{(1-\gamma)\beta} \psi y_H - E_H$ aL: $e^{-(1-\gamma)\beta} \frac{1-e^{-\gamma\beta}}{\gamma\beta} \psi y_L - E_L$	bL: $e^{-\gamma\beta} \frac{1-e^{-(1-\gamma)\beta}}{(1-\gamma)\beta} \psi y_L - E_L$ aL: $\frac{1-e^{-\gamma\beta}}{\gamma\beta} \psi y_L - E_L$

Define $\hat{\psi}_{aH,b}$ by $e^{-\gamma\beta} \frac{1-e^{-(1-\gamma)\beta}}{(1-\gamma)\beta} \psi y_H - E_H = e^{-\gamma\beta} \frac{1-e^{-(1-\gamma)\beta}}{(1-\gamma)\beta} \psi y_L - E_L$.

Define $\hat{\psi}_{aL,b}$ by $\frac{1-e^{-(1-\gamma)\beta}}{(1-\gamma)\beta} \psi y_H - E_H = e^{-\gamma\beta} \frac{1-e^{-(1-\gamma)\beta}}{(1-\gamma)\beta} \psi y_L - E_L$.

Define $\hat{\psi}_{bH,a}$ by $\frac{1-e^{-\gamma\beta}}{\gamma\beta} \psi y_H - E_H = e^{-(1-\gamma)\beta} \frac{1-e^{-\gamma\beta}}{\gamma\beta} \psi y_L - E_L$.

Define $\hat{\psi}_{bL,a}$ by $\frac{1-e^{-\gamma\beta}}{\gamma\beta} \psi y_H - E_H = \frac{1-e^{-\gamma\beta}}{\gamma\beta} \psi y_L - E_L$.

When $\gamma \geq \frac{1}{2}$, it can be verified that $\hat{\psi}_{aL,b} \leq \hat{\psi}_{bH,a} < \hat{\psi}_{bL,a} < \hat{\psi}_{aH,b}$. (1) We first prove for values of $\psi \in \left[0, \hat{\psi}_{aL,b}\right)$, the unique pure strategy Nash equilibrium is (aL, bL): group a choose low skill, group b choose low skill. Holding group b high skilled, group a choose to be low skilled, because $\psi < \hat{\psi}_{aL,b} < \hat{\psi}_{bH,a}$. Holding group b low skilled, group a choose to be low skilled, because $\psi < \hat{\psi}_{aL,b} < \hat{\psi}_{bL,a}$. Holding group a high skilled, group b choose to low skilled, because $\psi < \hat{\psi}_{aL,b} < \hat{\psi}_{aH,b}$. Holding group a low skilled, group b choose to be low skilled, because $\psi < \hat{\psi}_{aL,b}$. At the point $\psi = \hat{\psi}_{aL,b}$, the equilibrium can be (aL, bH) or (aL, bL).

(2) We prove for values of $\psi \in \left(\hat{\psi}_{aL,b}, \hat{\psi}_{bH,a}\right)$, the unique Nash pure strategy equilibrium is (aL, bH). Holding group b high skilled, group a choose to be low skilled, because $\psi < \hat{\psi}_{bH,a}$. Holding group b low skilled, group a choose to be low skilled, because $\psi < \hat{\psi}_{bH,a} < \hat{\psi}_{bL,a}$. Holding group a high skilled, group b choose to low skilled, because $\psi < \hat{\psi}_{bH,a} < \hat{\psi}_{aH,b}$. Holding group a low skilled, group b choose to be high skilled, because $\psi > \hat{\psi}_{aL,b}$. At the point $\psi = \hat{\psi}_{bH,a}$, the unique equilibrium is (aL, bH).

(3) We prove for values of $\psi \in \left(\hat{\psi}_{bH,a}, \hat{\psi}_{bL,a}\right)$, there is no pure strategy Nash equilibrium is (aL, bH). Holding group b high skilled, group a choose to be high skilled, because $\psi > \hat{\psi}_{bH,a}$. Holding group b low skilled, group a choose to be low skilled, because $\psi < \hat{\psi}_{bL,a}$. Holding group a high skilled, group b choose to low skilled, because $\psi < \hat{\psi}_{bL,a} < \hat{\psi}_{aH,b}$. Holding group a low skilled, group b choose to be high skilled, because $\psi > \hat{\psi}_{aL,b}$. At the point $\psi = \hat{\psi}_{bL,a}$, the unique equilibrium is (aH, bL).

(4) We prove for values of $\psi \in \left(\hat{\psi}_{bL,a}, \hat{\psi}_{aH,b}\right)$, there is a unique pure strategy Nash equilibrium (aH, bL). Holding group b high skilled, group a choose to be high skilled,

because $\psi > \hat{\psi}_{bH,a}$. Holding group b low skilled, group a choose to be high skilled, because $\psi > \hat{\psi}_{bL,a}$. Holding group a high skilled, group b choose to low skilled, because $\psi < \hat{\psi}_{aH,b}$. Holding group a low skilled, group b choose to be high skilled, because $\psi > \hat{\psi}_{aL,b}$. At the point $\psi = \hat{\psi}_{aH,b}$, the equilibrium can be (aH, bL) or (aH, bH).

(5) We prove for values of $\psi \in \left(\hat{\psi}_{aH,b}, 1\right)$, there is a unique pure strategy Nash equilibrium (aH, bH). Holding group b high skilled, group a choose to be high skilled. Holding group b low skilled, group a choose to be high skilled. Holding group a high skilled, group b choose to high skilled, because $\psi < \hat{\psi}_{aH,b}$. Holding group a low skilled, group b choose to be high skilled, because $\psi > \hat{\psi}_{aL,b}$.

(6) At last, notice that $\hat{\psi} = \frac{E_H - E_L}{y_H - y_L} \times \left(\frac{1-e^{-\beta}}{\beta}\right)^{-1}$, $\hat{\psi}_{bL,a} = \frac{E_H - E_L}{y_H - y_L} \times \left(\frac{1-e^{-\gamma\beta}}{\gamma\beta}\right)^{-1}$, and $\hat{\psi}_{aH,b} = \frac{E_H - E_L}{y_H - y_L} \times \left(e^{-\gamma\beta}\frac{1-e^{-(1-\gamma)\beta}}{(1-\gamma)\beta}\right)^{-1}$. It is straightforward to verify that $\hat{\psi}_{bL,a} < \hat{\psi} < \hat{\psi}_{aH,b}$ because $\left(\frac{1-e^{-\gamma\beta}}{\gamma\beta}\right)^{-1} < \left(\frac{1-e^{-\beta}}{\beta}\right)^{-1} < \left(e^{-\gamma\beta}\frac{1-e^{-(1-\gamma)\beta}}{(1-\gamma)\beta}\right)^{-1}$. Q.E.D.

Chapter 4

“Market Power and Efficiency in a Search Model” Revisited

Abstract

In a directed search context, Galenianos, Kircher and Virag (2011, IER) show how does firms' market power in a finite market lead to misallocation of workers, hence efficiency loss; they find that unemployment benefit could restore the efficiency, while the minimum wage can not. I suggest that it is important to make distinction between the qualitative existence of the market power and its quantitatively asymmetric impact when firms are heterogenous, and this distinction leads us to spot the distortion on allocation of surplus. Neither unemployment benefit nor minimum wage alone could correct the misallocation of labor and surplus at the same time, while a combination of both policies could. I further question the windfall nature of unemployment benefit, and show that if unemployment benefit is redistributively financed by the firms, misallocation of labor can never be rectified, while if it is financed by workers' wage tax, efficiency can result.

4.1 Introduction

Based on a wage posting context with search friction, Galenianos, Kircher and Virag (2011, IER) (hereafter GKV) argue that *"firms in a finite market*[1] *enjoy market power, in the sense that a single firm's action affects the equilibrium outcome of all agents"*; they show that with firms heterogenous in productivity and workers all identical, there is efficiency loss when the social planner's objective is to maximize the aggregate welfare. They further establish that unemployment benefit could restore the efficiency, while the minimum wage leads to further misallocation of labor.

There is an immediate inconsistency in their exposition: given their definition for the *market power* in a finite market, their result should have still been true in the particular case where firms do not differ in productivity; However, it is straightforward to observe that with firms identical in productivity, the market allocation of labor corresponds exactly to the social planner's allocation, hence there would be no longer efficiency loss in the sense of GKV (2011). What is the cause for this "disappearance" of firms' market power? I argue that the efficiency loss in their paper is in fact due to a compounded effect of two layers: (1) on the one hand, *qualitatively*, there is the classical notion of market power, which in our context is interpreted as individual firm's ability of using its posted wage to modify the labor supply curve; (2) on the other hand, *quantitatively*, there is what I term as the "asymmetry of market power" in the sense that the market power of firms with different productivity has asymmetrically different impact in magnitude to the allocation of workers. Both effects take place when firms are heterogenous in productivity, which is studied in GKV (2011); in case heterogeneity is absent, still present is the first effect, which contributes to the distortion on the allocation of surplus: firms tend to pay workers too less and grasp more surplus (or profit). And this distortion on surplus could be confirmed through decentralizing the social planner's problem. Our first result then simply shows that firms earn too much expected profits and the workers receive too few expected surplus and actual wage compared to the case absent of market power. As we will argue, confounding the two effects does not invalid their analysis on efficiency loss, but raises more concern for the policy maker.

GKV (2011) further show that introducing unemployment benefit could restore the efficiency; but introducing minimum wages exacerbates the misallocation of labor (related to asymmetry of market power), leading to further efficiency loss. I come up with several suggestions to complement their results. Firstly, I show that even introducing unemployment

[1] Here, finite market is referred to a market where the number of workers and firms are not infinite. We refer readers to GKV (2011) on a discussion of the significance of consideration of such a context.

benefit at the level which allows efficiency to be restored, workers' expected surplus is still lower than the socially optimal level. Secondly, although minimum wage detriments aggregate welfare by increasing misallocation of labor, there exists a feasible value of it which could fully correct the misallocation of surplus. In this sense, unemployment benefit is not necessarily a dominant choice over minimum wage. Thirdly, a proper level of combination of both policies a la GKV can correct both misallocations. I then argue further that direct comparison of the two policies is not appropriate, because between them lies another sharp difference - the source of funding. Notice that the unemployment benefit in GKV (2011) is introduced in a windfall manner. I suggest that it is more appropriate to consider a situation where this unemployment benefit is financed redistributively. I consider two situations: either the unemployment benefits are collected from the firms by unemployment tax or from the workers by wage tax. It is then shown that under the former taxation system, the efficiency in the sense of GKV (2011) does not obtain, but under the latter efficiency does obtain.

As a quick summary, we revisit GKV (2011) from a traditional Industrial Organization perspective. On one hand, taking both misallocation of labor and surplus into account necessitates us to rethink over the effectiveness of the policies involved; on the other hand, it allows us to have an impression on the possible phenomena which can appear after the policy is implemented; for example, if we believe that firms have market power and introduce the unemployment benefit in a windfall manner, then we expect to observe an increase in unemployment rate and wage gap, but we know that it is for the benefit of aggregate welfare. The consideration of such a finite labor market where firms have market power is of realistic importance. For example, Shimer (2007) mentions that if we segment labor market by the characteristics pair "occupation, geographical area", then a small number of vacancies may be potentially facing just a little above 10 job seekers per segment of the market.

The paper is organized as follows. In section 2.1, 2.2 and 2.3, I will recapitulate the original model with my remarks; in section 2.4, I state my arguments; in section 3, I consider the source of funding, and I will conclude in section 4.

4.2 The model

Consider an economy with two types of agents: the workers and the firms. We focus on the case with two workers and two firms, it allows us to capture most of the important results within a relatively simple context.

Workers are identical, and firms are ranked by their productivity y_j, with $y_1 \geq y_2$. We assume that the productivity of any firm is common knowledge. The wage posting game proceeds with the following timing:

1. Each firm posts a wage w_j.

2. Workers observe all the wages announcement $\mathbf{w} = (w_1, w_2)$, *and they choose which wage to apply to simultaneously.*

3. Firms collect applications. If more than one worker appears at one firm, this firm chooses one randomly; if the firm receives no application, then the post of this firm will end up vacant.

As in standard directed search literature, we search for the subgame perfect equilibria where firms use pure strategy on wage announcements and workers use mixed strategy in the job application subgame. We will also restrict attentions on symmetric equilibrium, where any worker applies to the same level of wage with identical probability, and any firm selects workers with identical probability. These assumptions help us introduce search friction in a natural way.

4.2.1 The market economy

This wage posting game is solved backwards by convention, and we start from the application subgame, where the workers, observing wage announcements $\{w_1, w_2\}$, choose which wage to apply to strategically.

Workers' application subgame. Name the two workers by A and B. Consider worker A's problem without loss of generality. Given an arbitrary wage pair $\{w_1, w_2\}$, his strategy is which firm to apply to. Define his probability of applying to firm j as α_j^A, then his probability of visiting the other firm is $1 - \alpha_j^A$. We proceed to find the matching probability for worker A conditional on his sending application to firm j.

Conditional on visiting firm j, the worker in question will be hired with probability 1 when the other worker did not apply to this firm j, and will be hired with probability $\frac{1}{2}$ when the other worker's application also arrived at this firm. The former event happens with probability $\left(1 - \alpha_j^B\right) \times 1$, and the latter event happens with probability $\alpha_j^B \times \frac{1}{2}$. In total, this worker's employment probability becomes

$$G\left(\alpha_j^B\right) = \frac{1}{2} \times \alpha_j^B + \left(1 - \alpha_j^B\right)$$

Analogously, worker B's employment probability becomes

$$G\left(\alpha_j^A\right) = \frac{1}{2} \times \alpha_j^A + \left(1 - \alpha_j^A\right)$$

A worker's expected payoff (or expected surplus) is the product of the wage he applied for and the probability of obtaining it: $U_j = G\left(\alpha_j^i\right) \times w_j$ where $i \in \{A, B\}$. Taking α_j^B, w_1 and w_2 as given, the worker 1's best response of visiting firm 1 is

$$\alpha_1^A = \begin{cases} 0 & \text{if } U_1 < U_2 \\ 1 & \text{if } U_1 > U_2 \\ (0,1) & \text{if } U_1 = U_2 \end{cases}$$

Only in case when the expected payoff from applying to both firms is equivalent, will the worker mix between these two firms. On the one hand, the payoff $G\left(\alpha_1^A\right) \times w_1$ the workers expect from applying to firm 1 should be no less than the payoff obtained from applying to firm 2; otherwise no worker will apply to firm 1. On the other hand, if payoff at firm 1 is strictly greater than the payoff at firm 2, no worker will apply to firm 2, hence the equality. We remark that when $U_1 = U_2$ holds, it defines workers' labor supply curve: $\frac{G\left(\alpha_1^A\right)}{G\left(\alpha_2^A\right)} = \frac{w_2}{w_1}$. It is straightforward to verify that the left-hand side of this equality is a decreasing function of α_1^A.

Firms' wage posting stage. We move to the firms' wage posting strategies. Take firm 1's problem as an example and the case of firm 2 can be solved in similar way. Firm 1 will be able to start production if at least one worker appears. The event that no worker appears happens with probability $\left(1 - \alpha_1^A\right)\left(1 - \alpha_1^B\right)$, hence at least one worker appears with probability $1 - \left(1 - \alpha_1^A\right)\left(1 - \alpha_1^B\right)$. A firm's payoff is simply

$$\left(1 - \left(1 - \alpha_1^A\right)\left(1 - \alpha_1^B\right)\right)(y_1 - w_1)$$

Intuitively, a high wage will induce more workers to apply, however, it leaves few ex post surplus to the firms. The dependence of the worker's application strategy on wage, i.e. a functional relationship between α and $\mathbf{w}$, is captured by

$$G\left(\alpha_1^B\right)\times w_1 = G\left(1-\alpha_1^B\right)\times w_2$$

for worker A, and

$$G\left(\alpha_1^A\right)\times w_1 = G\left(1-\alpha_1^A\right)\times w_2$$

for worker B, where the left- and right-hand side are respectively the expected payoff from applying to firm 1 and firm 2. *Since workers are identical, they use the symmetric strategy*: $\alpha_j^A = \alpha_j^B = \alpha_j$. The payoff is thus reduced to $\left(1-(1-\alpha_1)^2\right)(y_1 - w_1)$ for the firms, and $G(\alpha_1)\times w_1 = G(1-\alpha_1)\times w_2$ for workers.

We consider firm 1's optimization problem. Taking firm 2's strategy w_2, as well as the functional relationship between α_1 and w_1 as given, firm 1 maximizes the expected profit:

$$\begin{array}{cc} \max_{w_1} & \left(1-(1-\alpha_1)^2\right)\times(y_1-w_1) \\ \text{s.to} & \left(1-\frac{\alpha_1}{2}\right)\times w_1 = \left(1-\frac{\alpha_2}{2}\right)\times w_2 \end{array}$$

To solve this problem, we firstly substitute out w_1 according to the binding indifference constraint.

$$\left(1-(1-\alpha_1)^2\right)\times y_1 - \left(1-(1-\alpha_1)^2\right)\times\frac{1+\alpha_1}{2-\alpha_1}\times w_2$$

Differentiating with respect to α_1, we obtain the first order condition

$$2(1-\hat{\alpha}_1)\times y_1 - (1+2\hat{\alpha}_1)\times w_2 \;=\; 0$$

If we proceed analogously for firm 2, we have

$$2(1-\hat{\alpha}_2)\times y_2 - (1+2\hat{\alpha}_2)\times w_1 \;=\; 0$$

Moving the terms with w_1 and w_2 to the right-hand side respectively and make the ratio, we obtain

$$
\begin{aligned}
\frac{(1-\hat{\alpha}_1)\times y_1}{(1-\hat{\alpha}_2)\times y_2} &= \frac{(1+2\hat{\alpha}_1)\times w_2}{(1+2\hat{\alpha}_2)\times w_1} \\
&= \frac{(1+2\hat{\alpha}_1)\times(2-\hat{\alpha}_1)}{(1+2\hat{\alpha}_2)\times(2-\hat{\alpha}_2)}
\end{aligned}
$$

where the second equality comes from workers' labor supply curve. Rearrange to get

$$
\frac{y_1}{y_2} = \frac{(1+2\hat{\alpha}_1)\times(2-\hat{\alpha}_1)\times(1-\hat{\alpha}_2)}{(1+2\hat{\alpha}_2)\times(2-\hat{\alpha}_2)\times(1-\hat{\alpha}_1)}
$$

where $\hat{\alpha}_2 = 1-\hat{\alpha}_1$. The right-hand side hence is a function of $\hat{\alpha}_1$. It could be shown that the right-hand side is increasing, so that there is a unique solution to this polynomial. Term the solution as $\hat{\alpha}_1(y_1, y_2)$. According to the first order conditions, the equilibrium wages are then given by

$$
\hat{w}_1(y_1, y_2) = \frac{2(1-\hat{\alpha}_2(y_1,y_2))\times y_2}{(1+2\hat{\alpha}_2(y_1,y_2))}
$$

and

$$
\hat{w}_2(y_1, y_2) = \frac{2(1-\hat{\alpha}_1(y_1,y_2))\times y_1}{(1+2\hat{\alpha}_1(y_1,y_2))}
$$

Results (GKV) 1. When $y_1 > y_2$,
(1) The workers application strategies are such that $\hat{\alpha}_1 > \frac{1}{2} > \hat{\alpha}_2$
(2) The wages are such that $\hat{w}_1 > \hat{w}_2$
(3) The expected unemployment rate is $\frac{1}{2}\times\left(1-2\hat{\alpha}_1+2(\hat{\alpha}_1)^2\right)$

A high productivity firm in the equilibrium offers higher wage and attracts the workers to visit them more often than the low productivity firm.

Market payoff and market power

Given the wage announcements, workers are allowed to calculate the corresponding expected payoff at each firm which is a product of the actual wage and the probability of obtaining it. Based on all these calculated expected payoffs, workers decide which firm to apply to: for example, if the expected payoff from firm 1 is strictly lower than that provided by firm 2, then no workers will apply to firm 1.

To attract a worker, firm 1 should offer an expected payoff such that $\frac{2-\alpha_1}{2} \times w_1 \geq \frac{2-\alpha_2}{2} \times w_2$, analogously, firm 2 should offer an expected wage such that $\frac{2-\alpha_2}{2} \times w_2 \geq \frac{2-\alpha_1}{2} \times w_1$. The workers' supply curve is hence pinned down: $\frac{2-\alpha_1}{2} \times w_1 = \frac{2-\alpha_2}{2} \times w_2$. Then when we search for a mixed strategy equilibrium, each firm understands that the expected payoff offered by it must attain certain *reference level*, which is the expected payoff offered by the other firm. This reference level depends on the wages posted by both firms,[2] which implies that an individual firm's action has impact on workers' supply curve. In this sense, we say that firms have *market power*.

To see more exactly why market power is distortive, it is useful to introduce an aggregate variable U named *market payoff*, which is defined as the payoff that workers receive in the equilibrium of the subgame, or equivalently $U = \max\{U_1, U_2\}$. So that $\frac{2-\alpha_1}{2} \times w_1 = \frac{2-\alpha_2}{2} \times w_2 = U(w_1, w_2)$. By this equality, We can express α_1 (and analogously α_2) as a function of w_1 and w_2 as follows:

$$\alpha_1 = 2 \times \left(1 - \frac{U(w_1,w_2)}{w_1}\right)$$

Differentiating with respect to w_1, we obtain:

$$\frac{d\alpha_1}{dw_1} = \underset{>0}{\frac{\partial \alpha_1}{\partial w_1}} + \underset{<0}{\frac{\partial \alpha_1}{\partial U}} \times \underset{>0}{\frac{\partial U}{\partial w_1}}$$

In case $\frac{\partial U}{\partial w_j} = 0$, an individual firm's strategy has no impact on the aggregate variable U, and I show later that in this case the socially optimal allocation can result. Here, $\frac{\partial U}{\partial w_j} > 0$, so a distortive term $\frac{\partial \alpha_1}{\partial U} \times \frac{\partial U}{\partial w_1}$ is present; this term introduces a wedge in the marginal benefit for the firms, leading firms to post lower wages.

4.2.2 The social planner's problem

We move to the social planner's problem, and establish that the market allocation differs from the social planner's allocation. The social planner maximizes aggregate welfare of the economy, which coincides with the aggregate output due to risk neutrality. The objective is written as follows:

$$\begin{array}{ll} & \max_{\alpha_1,\alpha_2} \left(1-(1-\alpha_1)^2\right) y_1 + \left(1-(1-\alpha_2)^2\right) y_2 \\ s.to & \alpha_1 + \alpha_2 = 1 \end{array}$$

It could be shown that the social planner's allocation features

[2]From the labor supply curve, we observe that α_1 (hence α_2) is a function of both w_1 and w_2.

$$\frac{y_1}{y_2} = \frac{\alpha_1^p}{1-\alpha_1^p}$$

Result 2 (GKV): $\alpha_2^p < \hat{\alpha}_2 < \frac{1}{2} < \hat{\alpha}_1 < \alpha_1^p$

The result suggests that in the market economy workers apply to the high productivity firm less frequently. When market payoff depends on wages, an additional term will distort downwards marginal benefit, which renders firms' profits to be maximized at inferior levels of wage. When both firms reduce wages, the high productivity firm reduces wage disproportionately to a larger extent, which results in the misallocation of workers. There is also implication on the expected unemployment rate. Notice that the expected unemployment rate, $\frac{1}{2} \times \left((\alpha_1)^2 + (1-\alpha_1)^2 \right)$, is a parabola, decreasing for $\alpha_1 \in \left[0, \frac{1}{2}\right)$, minimized at $\alpha_1 = \frac{1}{2}$, and increasing for $\alpha_1 \in \left(\frac{1}{2}, 1\right]$. In the market economy, this rate is simply lower compared to the social planner's problem, and it is because workers apply to the low productivity firm too often. Hence, there is a conflicting objective between minimizing the expected unemployment rate and maximizing the aggregate welfare.

4.2.3 Policy implications

GKV (2011) then show that: *Unemployment benefit can restore efficiency, while minimum wage can not.* Let me give more details on how they obtain the result.

Minimum wage. When minimum wage policy is effective, the low productivity firm's wage is raised mechanically to a higher level. This leads firm 2 to attract even more frequent application from workers, hence further exacerbates the problem of the misallocation of labor, pushing the economy further away from efficiency.

Technically, suppose now that the government intervenes by imposing a minimum wage level $\underline{w_2}$. Notice that in case $\underline{w_2} < \hat{w}_2$, the minimum wage policy simply affects no firm and has no impact. We hence focus on the case where we have $\underline{w_2} \geq \hat{w}_2$, so that when this minimum wage policy is enforced, the equilibrium value of wage announced by firm 2 will be simply $\underline{w_2}$ instead of $\hat{w}_2$. When this minimum wage is effective for firm 2, firm 2 is not able to maximize profit. From workers' indifference equation $\frac{2-\alpha_1}{2} \times w_1 = \frac{2-\alpha_2}{2} \times \underline{w_2}$, we express simply workers' application strategy α_1 as a function of $\left\{w_1, \underline{w_2}\right\}$. Let firm 1 maximize with respect to w_1, from its first order condition $2\left(1-\underline{\hat{\alpha}_1}\right) \times y_1 - \left(1+2\underline{\hat{\alpha}_1}\right) \times \underline{w_2} = 0$, we notice that $\underline{\hat{\alpha}_1}$ is simply inferior to $\hat{\alpha}_1$. This implies that the misallocation is even more severe compared to the case where minimum wage is not introduced.

GKV (2011) remark that the original economy, where minimum wage is imposed to firm 2, is equivalent to an alternative economy where firm 2 has a productivity level $\tilde{y}_2 > y_2$. By this parallelization, they find that minimum wage makes the firm 2 hire even more often so that the inefficiency becomes more severe. Intuitively, when the productivity gap between two firms reduces, the impact from the asymmetry of market power decreases, then both α_1 and α_2 tend to be closer to $\frac{1}{2}$.

Unemployment benefit. GKV (2011) then introduces unemployment benefit in a windfall manner. Since unemployment benefit raises workers' outside option, workers are willing to take more "risk". Although the event of being not employed occurs more often in the high productivity firm, workers could receive the unemployment benefit at a more frequent basis compared to applying to the low productivity firm. So as the benefit increases, workers shift gradually their visiting frequency to the high productivity firm. GKV (2011) establish that under this particular context, such an unemployment benefit could correct misallocation of labor.

Technically, with unemployment benefits, the model could be rewritten as follows:

$$\begin{array}{rc} \max_{w_1} & \left(1-(1-\alpha_1)^2\right) \times (y_1 - w_1) \\ \text{s.to} & \left(1-\frac{\alpha_1}{2}\right) \times (w_1 - b) = \left(1-\frac{\alpha_2}{2}\right) \times (w_2 - b) \end{array}$$

for firm 1, and

$$\begin{array}{rc} \max_{w_2} & \left(1-(1-\alpha_2)^2\right) \times (y_2 - w_2) \\ \text{s.to} & \left(1-\frac{\alpha_1}{2}\right) \times (w_1 - b) = \left(1-\frac{\alpha_2}{2}\right) \times (w_2 - b) \end{array}$$

for firm 2. By making a change of variable from $W_j = w_j - b$, it could be shown that this economy is equivalent to an economy where y_1 becomes $y_1 - b$, and y_2 becomes $y_2 - b$. Thus in order to determine whether introducing b could raise efficiency, it suffices to examine whether the following expression

$$\frac{y_1 - b}{y_2 - b} = \frac{\left(1+2\hat{\alpha}_1^b\right) \times \left(2-\hat{\alpha}_1^b\right) \times \left(1-\hat{\alpha}_2^b\right)}{\left(1+2\hat{\alpha}_2^b\right) \times \left(2-\hat{\alpha}_2^b\right) \times \left(1-\hat{\alpha}_1^b\right)}$$

leads to a higher level of $\hat{\alpha}_1^b$ when b increases. The answer is affirmative, since the left-hand side is increasing in b, and the right-hand side is also increasing in $\hat{\alpha}_1^b$. GKV (2011) concludes that there exists an unemployment benefit level such that the corresponding $\hat{\alpha}_1^b$

could be brought back to the efficient level $\hat{\alpha}_1^p$, so that it is a more effective policy under this particular economy.

An accompanying result is that with b introduced, the wage gap $w_1 - w_2$ actually increases. Hence, if in reality it is believed that firms have market power, then when b is introduced we may observe both increase in wage inequality and increase in the unemployment rate controlling for workers' ability.

4.2.4 Critics

On the source of inefficiency

GKV (2011) note that *"firms in a finite market enjoy market power, in the sense that a single firm's action affects the equilibrium outcome of all agents"*. According to this definition, market power should exist in any market with finite number of workers and firms. However, it is immediate to observe that, in the original model, if we impose $y_1 = y_2$, i.e. if firms have identical productivity, then the optimal solution $\hat{\alpha}_1$ in the market economy coincides with the choice α_1^p in the social planner's problem: $\hat{\alpha}_1 = \alpha_1^p = \frac{1}{2}$. Consequently, according to the interpretation of GKV (2011) there will not be any efficiency loss because there is no loss of the aggregate output in the market economy.

Does it mean that the effect of market power vanishes in case $y_1 = y_2$? Not really. Recall the interpretation of GKV (2011) on their model with $y_1 > y_2$: *"the high productivity firm will respond to market power by decreasing its wage by a larger amount, which leads to the misallocation of workers"*.[3] Then in the extreme case with $y_1 = y_2$, the result $\hat{\alpha}_1 = \alpha_1^p = \frac{1}{2}$ is simply due to the fact that firms respond to market power by decreasing their wages to the same magnitude, such that the allocation on the application of workers is not disturbed. The difference is plainly on what I term as *the asymmetry of the market power*. That is to say, although both firms lower their wages when taking into consideration their market power, different productivity leads to disproportionately different amounts of wages reduction; so that workers' application decision is distorted due to the quantitatively asymmetric impact from the reduction of wages by firms with different productivities. And more importantly, although in case of identical productivity this asymmetry of market power is not in action, the market power is still qualitatively present: in a finite market, *"firms face less competition and they post lower wages than they would if they take their market power into consideration"*;[4]

[3] *See GKV (2011) Page 10.*
[4] See GKV (2011) page 9

I then suggest that even when the allocation of labor is not distorted, the allocation of surplus can be distorted: firms can enjoy higher expected surplus than they would have if they do not take their market power into consideration.

To show that in the market economy firms enjoy higher expected surplus and workers suffer from lower expected surplus, it suffices to decentralize social planner's problem, and then make the comparison. Formally, according to section 2.1.1., in the decentralized social planner's problem we could simply neglect the firms' taking into account of their market power by forcing the aggregate variable market payoff U invariant to any individual firms' price variation, i.e. $\frac{\partial U}{\partial w_j} = 0$. For example, firm 1's problem now becomes:

$$\begin{array}{ll} \max_{w_1} & \left(1-(1-\alpha_1)^2\right) \times (y_1 - w_1) \\ \text{s.to} & \left(1-\frac{\alpha_1}{2}\right) \times w_1 = U \end{array}$$

where U is invariant to w_1 and w_2. Solving the first order condition, we obtain that $U = \left(1-\alpha_1^D\right) y_1$. Analogously, we obtain $U = \left(1-\alpha_2^D\right) y_2$ from firm 2's optimization problem. By these two equalities we obtain $\frac{y_1}{y_2} = \frac{1-\alpha_2^D}{1-\alpha_1^D}$, which pins down the level of α_1^D. We verify that the so obtained α_1^D coincides with social planner's allocation $\alpha_1^D = \alpha_1^p$. We have then the following proposition:

Proposition 1. *Compared to the efficient level, in the market economy (1) firms enjoy higher expected surplus and workers suffer from lower expected surplus; (2) workers suffer from lower actual wages.*

Proof. In the Appendix.

As remarked above, when the market payoff depends on wages, a distortive term in the firm's marginal benefit arises, which leads firms to post lower wages when they maximize profit. This suggests that the surplus is misallocated towards the firms, and the low wages imply low market payoff. In total there is deadweight loss and the aggregate output is lower. It is also important to point out that when market power is present, the equilibrium functional relationship between the wage w_j and the application strategy α_j is different in problem of the market economy compared to that of the decentralized economy. So that even if the allocation of labor is corrected to the socially optimal level, the misallocation of surplus persists - a result summarized in the following corollary.

Corollary. *Even if we replace $\hat{\alpha}_1$ by α_1^p in the market economy, the resulted workers' expected surplus is lower than its efficient level.*

On policy

Policy 1: Unemployment benefit. In the section 2.3, GKV (2011) show that unemployment benefit could restore the efficiency: by introducing a proper level of b, α_1^b could be rectified to the efficient level α_1^p; so that the aggregate output where unemployment benefit is introduced could reach the socially optimal level.

$$\begin{aligned} & \left(1-\left(1-\alpha_1^b\right)^2\right) y_1+\left(1-\left(1-\alpha_2^b\right)^2\right) y_2 \\ = & \left(1-\left(1-\alpha_1^p\right)^2\right) y_1+\left(1-\left(1-\alpha_2^p\right)^2\right) y_2 \end{aligned}$$

However, in the section 2.4.1 we see that this merely corrects the misallocation of workers, not the misallocation of surplus. Now define the level of b at which the efficiency is restored in the sense of GKV (2011) as b^A. We have the following result showing that even taking into account this b^A, workers' expected surplus is inferior to its efficient level. Define the workers' expected surplus in the market economy with unemployment benefit as $U\left(b^A\right)$, we have the following proposition:

Proposition 2. *In the market economy augmented with unemployment benefit b^A, the workers' expected surplus $U\left(b^A\right)$ is strictly lower than the efficient level U^p.*

Proof. In the Appendix.

As a result, if the allocation of labor and of surplus are both of concern, the level b required to correct the misallocation of labor does not at the same time bring workers' surplus back to its efficient level. If we further increase b beyond b^A, we are able to increase further the workers' surplus; however, the misallocation of labor will be distorted again, but to such an extent that workers apply too often (more than the efficient level) to the high productivity firm, where they are employed at less frequent basis. Hence, if b is the only instrument at the policy maker's disposal, both misallocations can not be addressed at the same time. We have the following comparative statics to see how the productivity gap has impact on firms' market power.

Proposition 3 (comparative statics). *The level of b^A increases with the productivity gap y_1-y_2.*

Proof. In the Appendix.

The result suggests that as the productivity gap becomes larger, the level of b required to correct the misallocation of labor increases. Intuitively, as the productivity gap increases, the reduction of wage from the high productivity firm becomes more and more important relatively to that from the low productivity firm, which increasingly worsens the misallocation of labor.

Policy 2: Minimum wage. In GKV (2011), it is shown that minimum wage has the effect of discouraging the workers from visiting the high productivity firm, leading to misallocation of labor. By this, GKV (2011) concludes that minimum wage is a less effective political instrument compared to unemployment benefit. This result should be taken with caution. Indeed, although minimum wage further distorts the allocation of labor, it increases workers' expected surplus, hence can correct the misallocation of surplus:

Proposition 4. *There exists a level of $\underline{w}^c \in (\hat{w}_2, y_2)$ such that the misallocation of surplus can be completely corrected.*

Proof. In the Appendix.

A sharp difference between the minimum wage and the kind of unemployment benefit introduced by GKV (2011) is that the former does not cost government's budget - in their context, the unemployment benefit is essentially regarded as coming from external resources - a windfall. Since Proposition 2 implies that the level of b required to correct the misallocation of surplus is greater than that required for the misallocation of labor, minimum wage is more effective in correcting the surplus distortion. Before entering into the next section, we suggest the natural result that a combination of unemployment benefit and minimum wage can correct the misallocation of labor and surplus at the same time:

Corollary. *Introducing a proper level of unemployment benefit, together with minimum wage, can correct both the misallocation of labor and misallocation of surplus.*

Proof. In the Appendix.

4.3 Further discussions

4.3.1 On funding source of unemployment benefit

In this section, we focus exclusively on the misallocation of labor. We want to examine whether GKV (2011)'s result on unemployment benefit could still hold if the funding source is not of a windfall manner. Given that firms enjoy extra profit from their market power, a natural starting point is a type of redistributive economy where the firms pay for the unemployment consequence caused by themselves.

We name now the taxation system in GKV(2011) as system A. We then follow Geromichalos (2014) to study one type of funding for unemployment benefit which we call taxation system B: given an unemployment benefit level b, each firm pays to each worker who applied to it but did not get employed an equal amount of unemployment benefit. We have the following program for firm 1 (firm 2's problem is analogous),

$$\begin{array}{ll} \max_{w_1} & \left(1-(1-\alpha_1)^2\right) \times (y_1 - w_1) - (\alpha_1)^2 b \\ \text{s.to} & \frac{2-\alpha_1}{2} \times (w_1 - b) = \frac{2-\alpha_2}{2} \times (w_2 - b) \end{array}$$

where α_1^2 represents the probability with which firm 1 has to pay the workers who applied to it but ended up not being selected by it. And We have the following proposition for the results:

Proposition 5. *In system B, there exists no unemployment benefit level b^B such that the efficiency is restored in the sense of GKV(2011).*

Proof. In the Appendix.

If we compare system B with the economy with no unemployment benefit, the introduction of the term $\left(\alpha_j\right)^2 b$ on one hand increases the marginal cost of firms, giving firms incentive to lower wage; on the other hand, workers demand higher wages due to the raised outside option b. Which effect dominates? My analysis shows that, the high productivity firm further decreases its wage hence the former effect dominates, while the low productivity firm increases its wage, hence the latter effect dominates. Recall that without unemployment benefit, both firms actually post too low wages, we can conclude that introducing this taxation system does not restore the efficiency caused by misallocation of labor.

We now propose another source of funding for the unemployment benefit: workers' wage taxation. We show that under this system the misallocation of labor could be corrected. Suppose workers pay a tax T when they get matched, and they will receive an amount $\eta \times T$ whenever they end up unemployed, where η is a scalar which keeps the planner's budget non-negative. When a worker applies to firm 1, her expected payoff is $\left(1-\frac{\alpha_1}{2}\right) \times (w_1 - T) + \frac{\alpha_1}{2}\eta T$; since she applies to firm 1 with probability α_1, her contribution to the tax system when applying to firm 1 is: $\alpha_1 \times \left[\left(1-\frac{\alpha_1}{2}\right) \times (-T) + \frac{\alpha_1}{2}\eta T\right]$. When she applies to firm 2, she expects $\left(1-\frac{\alpha_2}{2}\right)\times(w_2-T)+\frac{\alpha_2}{2}\eta T$; since she applies to firm 2 with probability α_2, her corresponding contribution to the tax system is: $\alpha_2\times\left[\left(1-\frac{\alpha_2}{2}\right)\times(-T)+\frac{\alpha_2}{2}\eta T\right]$. Since we require that η is such that the planner does not have any contribution, the system should be self-financing:

$$\alpha_1\times\left[\left(1-\frac{\alpha_1}{2}\right)\times(-T)+\frac{\alpha_1}{2}\eta T\right]+\alpha_2\times\left[\left(1-\frac{\alpha_2}{2}\right)\times(-T)+\frac{\alpha_2}{2}\eta T\right]\geq 0$$

Assuming T to be positive, it could be cancelled out. Rearranging, we obtain: $\eta\geq\frac{2}{(\alpha_1^2+\alpha_2^2)}-1$. Since $\alpha_1^2+\alpha_2^2<(\alpha_1+\alpha_2)^2=1$, we have $\eta\geq\frac{2}{(\alpha_1^2+\alpha_2^2)}-1>2-1=1$. So that one unit of wage taxation can be redistributively translated into $\eta>1$ units of unemployment benefit. The important implication is that the planner could simply adjust the amount of unemployment benefit from ηT to T, by which the planner can gain some extra budget and there is always a T such that the misallocation of labor is corrected. To see how, we replace ηT by T, and since workers are indifferent from applying to any firm, we require $\left(1-\frac{\alpha_1}{2}\right)\times(w_1-T)+\frac{\alpha_1}{2}T=\left(1-\frac{\alpha_2}{2}\right)\times(w_2-T)+\frac{\alpha_2}{2}T$, which could be rewritten as $\left(1-\frac{\alpha_1}{2}\right)\times(w_1-2T)=\left(1-\frac{\alpha_2}{2}\right)\times(w_2-2T)$. And the firm 1's problem (firm 2's problem is similar) becomes:

$$\begin{array}{cc}\max_{w_1} & \left(1-(1-\alpha_1)^2\right)\times(y_1-w_1)\\ \text{s.to} & \frac{2-\alpha_1}{2}\times(w_1-2T)=\frac{2-\alpha_2}{2}\times(w_2-2T)\end{array}$$

Replacing $2T$ by b, GKV (2011) show that there is always a b (hence T) such that the misallocation of labor is corrected. As a quick summary, we have shown that the redistributive wage taxation system is superior compared to the redistributive unemployment benefit system where the firms pay the bills, because the former allows the misallocation of labor to be rectified, and may allow the government to earn extra budget. In the following section, we consider another type of policy intervention: imposing wage to be a fixed proportion of productivity.

4.3.2 Imposing wage to be a fixed proportion of productivity

Consider the case where wages are forced to represent a fixed proportion of productivity. Specifically, workers receive a fixed part of production βy_1 from firm 1 or βy_2 from firm 2, and firm 1 receives $(1-\beta)y_1$ and firm 2 receives $(1-\beta)y_2$. Workers know which firm is of which level of productivity, and they are still able to choose where to apply. The timing is as previous. We are interested in the case where workers are indifferent from applying to high and low productivity firms where they anticipate receiving βy_1 and βy_2 respectively, and the indifference condition becomes

$$\frac{2-\alpha_1^{\mathrm{NB}}}{2}\times\beta y_1=\frac{2-\alpha_2^{\mathrm{NB}}}{2}\times\beta y_2$$

Notice that α_1^{NB} is determined by y_1 and y_2, and does not depend on β due to risk neutrality.[5] Firstly, we are interested in the comparison of α_1^{NB} with the socially optimal level α_1^p which is determined by $\frac{y_1}{y_2}=\frac{\alpha_1^p}{1-\alpha_1^p}$, and the market level $\hat{\alpha}_1$ which is determined by $\frac{y_1}{y_2}=\frac{(1+2\hat{\alpha}_1)\times(2-\hat{\alpha}_1)\times(1-\hat{\alpha}_2)}{(1+2\hat{\alpha}_2)\times(2-\hat{\alpha}_2)\times(1-\hat{\alpha}_1)}$. It could be shown that we have

$$\hat{\alpha}_1<\alpha_1^p<\alpha_1^{\mathrm{NB}}$$

Hence, workers apply to the high productivity firm too often with fixed sharing rule. Secondly, we are interested to see how this rule affects the allocation of surplus. For this sake, we define β such that workers' expected surplus with bargaining achieves the socially optimal level U^p. Accordingly, β^{NB} is determined by $\frac{2-\alpha_1^{\mathrm{NB}}(y_1,y_2)}{2}\times\beta^{\mathrm{NB}}y_1=\frac{2-\alpha_2^{\mathrm{NB}}(y_1,y_2)}{2}\times\beta^{\mathrm{NB}}y_2=U^p$. We have the following claim summarizing the main results.

Claim. With fixed sharing rule β^{NB} such that workers' expected payoff is equal to U^p,

(1) the high (low) productivity firm offers higher (lower) wage compared to the case where wages are posted.

(2) both the high and low productivity firm earn lower profit compared to the case where wages are posted.

Proof. In the Appendix.

I compare this policy experiment with minimum wage. Recall that introducing minimum wage reduces the wage gap, and leads the workers to apply to the low productivity firm more often compared to the socially optimal case; With the fixed sharing rule, which introduces a perfect correlation between the wage of high and low productivity firms, the result is reversed: the wage gap widens, and the workers apply to the high productivity firm too often.

4.4 Conclusion

In this paper, I revisit GKV (2011), and emphasize the distinction between *the qualitative existence of market power* and *its quantitatively asymmetric impact due to productivity heterogeneity*. Making this distinction allows me to spot the misallocation of surplus, apart

[5]The condition of $y_1<2y_2$ should be imposed to make sure that $0<\alpha_1^{\mathrm{NB}}<1$.

from the misallocation of labor which is central in their context. I reexamine unemployment benefit and minimum wage policy on both misallocations. I then question the windfall nature of the unemployment benefit, and propose to study the case where this benefit is financed redistributively. If unemployment benefit is financed by the firms, efficiency can never result, but if it is financed by workers' wage taxes, efficiency can obtain.

To how much extent are the results robust? My preliminary analysis shows that all the results could hold true when we have finitely many firms and workers while limiting the degree of productivity heterogeneity to be two. At last, it may be argued that in GKV (2011), the misallocation of surplus is not crucial in the sense that the efficient level of aggregate welfare can be achieved as long as there is no misallocation of labor. The underlying reason for this particular result is the over-simplified functional form for the surplus of firms and workers. For example, it is straightforward to check that adding risk aversion will make the wages appear in the aggregate welfare, which makes the difference. We consider it as the direction for the next step of the current work.

4.5 Appendix

Proposition 1. *Compared to the efficient level, in the market economy (1) firms enjoy higher expected surplus and workers suffer from lower expected surplus; (2) workers suffer from lower actual wages.*

Proof. (1) Firstly, we prove that workers suffer from lower expected surplus. In the market economy, the level of market payoff is simply

$$
\begin{aligned}
U^M &= \tfrac{2-\hat{\alpha}_1}{2} \times \overbrace{\frac{2\left(1-\hat{\alpha}_2\right) \times y_2}{\left(1+2\hat{\alpha}_2\right)}}^{\hat{w}_1} \\
&= \tfrac{1+\hat{\alpha}_2}{(1+2\hat{\alpha}_2)} \times \left(1-\hat{\alpha}_2\right) \times y_2
\end{aligned}
$$

While the level of market utility in the planner's economy is simply

$$
U^p = \left(1-\alpha_2^p\right) \times y_2
$$

Since $\frac{1+\hat{\alpha}_2}{(1+2\hat{\alpha}_2)} < 1$, and $\left(1-\hat{\alpha}_2\right) < \left(1-\alpha_2^p\right)$, we have the result $U^M < U^p$ for the workers.

Seconly, we prove that firms enjoy higher surplus. In the market economy firm 2's expected surplus, after some arrangement, is

$$\pi^M_{firm2} = \left(1-(1-\hat{\alpha}_2)^2\right)y_2 - \frac{1+\hat{\alpha}_2}{1+2\hat{\alpha}_2}\times 2\hat{\alpha}_2(1-\hat{\alpha}_2)y_2$$

While in the decentralized economy, firm 2's expected surplus, after some arrangement becomes

$$\pi^D_{firm2} = \left(1-\left(1-\alpha_2^p\right)^2\right)y_2 - 2\alpha_2^p\times\left(1-\alpha_2^p\right)y_2$$

Since $\hat{\alpha}_2 > \alpha_2^p$, $\frac{1+\hat{\alpha}_2}{1+2\hat{\alpha}_2} < 1$ and π^D_{firm2} is increasing in α_2^p, we have

$$\begin{aligned} & \pi^M_{firm2} \\ > & \left(1-(1-\hat{\alpha}_2)^2\right)y_2 - 2\hat{\alpha}_2(1-\hat{\alpha}_2)y_2 \\ > & \pi^D_{firm2} \end{aligned}$$

As for firm 1, its surplus in the market economy is

$$\pi^M_{firm1} = \left(1-(1-\hat{\alpha}_1)^2\right)(y_1-\hat{w}_1)$$

where $\hat{\alpha}_1$ and $\hat{w}_1$ should satisfy the workers' indifference condition which can be rearranged to $\hat{\alpha}_1 = \frac{2\hat{w}_1-\hat{w}_2}{\hat{w}_1+\hat{w}_2}$. We notice that $\hat{\alpha}_1$ is increasing in $\hat{w}_1$ and decreasing in $\hat{w}_2$ according to this relationship.

While in the decentralized economy, firm 1's expected surplus is

$$\pi^D_{firm1} = \left(1-\left(1-\alpha_1^p\right)^2\right)\left(y_1-w_1^p\right)$$

where α_1^p and w_1^p should also satisfy workers' indifference condition which can be rearranged as $\alpha_1^p = \frac{2w_1^p-w_2^p}{w_1^p+w_2^p}$. Again, α_1^p is increasing in w_1^p and decreasing in w_2^p. We have

$$\begin{aligned} \pi^M_{firm1} = & \left(1-(1-\hat{\alpha}_1(\hat{w}_1,\hat{w}_2))^2\right)(y_1-\hat{w}_1) \\ > & \left(1-\left(1-\hat{\alpha}_1\left(w_1^p,\hat{w}_2\right)\right)^2\right)\left(y_1-w_1^p\right) \\ > & \left(1-\left(1-\hat{\alpha}_1\left(w_1^p,w_2^p\right)\right)^2\right)\left(y_1-w_1^p\right) \end{aligned}$$

where the first inequality is due to revealed preference, and the second inequality is due to the fact that $\hat{w}_2 < w_2^p$, a result which we are going to prove in the part 2 of the proof.

(2) Firstly, we prove $\hat{w}_1 < w_1^p$. We have $\hat{w}_1 = \frac{2(1-\hat{\alpha}_2)}{(1+2\hat{\alpha}_2)}\times y_2$, and $w_1^p = \frac{2\left(1-\alpha_2^p\right)}{1+\alpha_2^p}\times y_2$. Then we know that

$$\begin{aligned}\hat{w}_1 &= \tfrac{2(1-\hat{\alpha}_2)}{(1+2\hat{\alpha}_2)} \times y_2 \\ &< \tfrac{2(1-\hat{\alpha}_2)}{(1+\hat{\alpha}_2)} \times y_2 \\ &< \tfrac{2\left(1-\alpha_2^p\right)}{\left(1+\alpha_2^p\right)} \times y_2 = w_1^p\end{aligned}$$

where the second inequality is due to the fact that $\frac{1-x}{1+x}$ is a decreasing function and $\hat{\alpha}_2 > \alpha_2^p$.

Secondly, we prove $\hat{w}_2 < w_2^p$. We have $\hat{w}_2 = \frac{2(1-\hat{\alpha}_1)}{(1+2\hat{\alpha}_1)} \times y_1$, and $w_2^p = \frac{2\left(1-\alpha_2^p\right)}{2-\alpha_2^p} \times y_2$. Then we know that

$$\begin{aligned}\hat{w}_2 &= \tfrac{2(1-\hat{\alpha}_1)}{(1+2\hat{\alpha}_1)} \times y_1 \\ &= \tfrac{(1+\hat{\alpha}_2)}{(1+2\hat{\alpha}_2)} \tfrac{2(1-\hat{\alpha}_2)}{(2-\hat{\alpha}_2)} \times y_2 \\ &< \tfrac{2\left(1-\alpha_2^p\right)}{\left(2-\alpha_2^p\right)} \times y_2 = w_2^p\end{aligned}$$

where the second equality is obtained by the condition which determines $\hat{\alpha}_1$, i.e. $\frac{y_1}{y_2} = \frac{(1+2\hat{\alpha}_1)\times(2-\hat{\alpha}_1)\times(1-\hat{\alpha}_2)}{(1+2\hat{\alpha}_2)\times(2-\hat{\alpha}_2)\times(1-\hat{\alpha}_1)}$; the inequality is obtained because $\frac{(1+\hat{\alpha}_2)}{(1+2\hat{\alpha}_2)} < 1$ and $\frac{(1-\hat{\alpha}_2)}{(2-\hat{\alpha}_2)}$ is a decreasing function of $\hat{\alpha}_2$. **Q.E.D.**

Proposition 2. *In the market economy augmented with unemployment benefit b^A, the workers' expected surplus $U\left(b^A\right)$ is strictly lower than the efficient level U^p.*

Proof. Recall that b^A is determined by

$$\frac{y_1 - b^A}{y_2 - b^A} = \frac{\left(1+2\alpha_1^p\right) \times \left(2-\alpha_1^p\right) \times \left(1-\alpha_2^p\right)}{\left(1+2\alpha_2^p\right) \times \left(2-\alpha_2^p\right) \times \left(1-\alpha_1^p\right)}$$

We obtain then b^A as a function of α_1^p: $b^A = y_2 - \frac{y_1 - y_2}{\frac{\left(1+2\alpha_1^p\right)\times\left(2-\alpha_1^p\right)\times\left(1-\alpha_2^p\right)}{\left(1+2\alpha_2^p\right)\times\left(2-\alpha_2^p\right)\times\left(1-\alpha_1^p\right)} - 1} < y_2 < y_1$. Defining $K\left(\alpha_1^p\right) = \frac{\left(1+2\alpha_1^p\right)\times\left(2-\alpha_1^p\right)\times\left(1-\alpha_2^p\right)}{\left(1+2\alpha_2^p\right)\times\left(2-\alpha_2^p\right)\times\left(1-\alpha_1^p\right)}$, we have $b^A = y_2 - \frac{y_1-y_2}{K\left(\alpha_1^p\right)-1}$.

We start by calculating $U\left(b^A\right)$. By definition, we know that

$$\begin{aligned}U\left(b^A\right) &= \left(1 - \tfrac{\alpha_2^p}{2}\right) \hat{w}_2\left(b^A\right) + \left[1 - \left(1 - \tfrac{\alpha_2^p}{2}\right)\right] \times b^A \\ &= \left(1 - \tfrac{\alpha_2^p}{2}\right) \times \left(\hat{w}_2\left(b^A\right) - b^A\right) + b^A\end{aligned}$$

where $\hat{w}_2\left(b^A\right)$ is the equilibrium wage when b^A is introduced, and more precisely $\hat{w}_2\left(b^A\right)$ is the solution of the following program:

$$\begin{array}{cc} \max_{w_2} & \left(1-(1-\alpha_2)^2\right)\times(y_2-w_2) \\ \text{s.to} & \left(1-\frac{\alpha_1}{2}\right)\times(w_1-b)=\left(1-\frac{\alpha_2}{2}\right)\times(w_2-b) \end{array}$$

This program gives $\hat{w}_2\left(b^A\right)-b^A=\frac{2\left(1-\alpha_1^p\right)}{\left(1+2\alpha_1^p\right)}\times\left(y_1-b^A\right)$. We then have

$$\begin{aligned} U\left(b^A\right) &= \left(1-\frac{\alpha_2^p}{2}\right)\times\frac{2\left(1-\alpha_1^p\right)}{\left(1+2\alpha_1^p\right)}\times\left(y_1-b^A\right)+b^A \\ &= \left(1+\alpha_1^p\right)\times\frac{\left(1-\alpha_1^p\right)}{\left(1+2\alpha_1^p\right)}\times\left(y_1-b^A\right)+b^A \end{aligned}$$

We now make use of the expression of b^A which we obtained in the beginning of the proof, and we have

$$U\left(b^A\left(\alpha_1^p\right)\right) = \frac{\frac{1-\left(\alpha_1^p\right)^2}{1+2\alpha_1^p}\times K\left(\alpha_1^p\right)-1}{K\left(\alpha_1^p\right)-1}\times(y_1-y_2)+y_2$$

At the same time, we know that $U^p=\left(1-\alpha_1^p\right)y_1$. We could do the following arrangement:

$$\begin{aligned} U^p &= \left(1-\alpha_1^p\right)(y_1-y_2)+\left(1-\alpha_1^p\right)y_2 \\ &= \left(1-\alpha_1^p-\alpha_1^p\frac{y_2}{y_1-y_2}\right)(y_1-y_2)+y_2 \\ &= \left(1-\alpha_1^p\frac{1}{1-(y_2/y_1)}\right)(y_1-y_2)+y_2 \end{aligned}$$

Since $\frac{y_2}{y_1}=\frac{\left(1-\alpha_1^p\right)}{\alpha_1^p}$, we could finally obtain

$$\begin{aligned} U^p &= \left(1-\alpha_1^p\frac{1}{1-\frac{\left(1-\alpha_1^p\right)}{\alpha_1^p}}\right)(y_1-y_2)+y_2 \\ &= \left(1-\frac{\left(\alpha_1^p\right)^2}{2\alpha_1^p-1}\right)(y_1-y_2)+y_2 \end{aligned}$$

Then it suffices to compare the coefficient in front of (y_1-y_2) in $U\left(b^A\left(\alpha_1^p\right)\right)$ and U^p respectively. Numerically we are able to verify that $U\left(b^A\left(\alpha_1^p\right)\right)<U^p$. As shown in the following graph, the $X-axis$ represents α_1^p, the curve below corresponds to $U\left(b^A\left(\alpha_1^p\right)\right)$, and the curve above corresponds to U^p. **Q.E.D.**

Proposition 3 (comparative statics). *The level of b^A increases with the productivity gap y_1-y_2.*

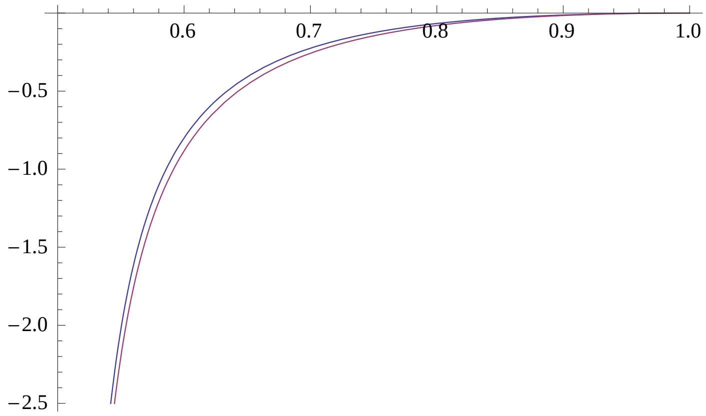

Fig. 4.1 Comparison between U^p and $U\left(b^A\left(\alpha_1^p\right)\right)$

Proof. Recall that b^A is such that

$$\frac{y_1 - b^A}{y_2 - b^A} = \frac{\left(1+2\alpha_1^p\right)\times\left(2-\alpha_1^p\right)\times\left(1-\alpha_2^p\right)}{\left(1+2\alpha_2^p\right)\times\left(2-\alpha_2^p\right)\times\left(1-\alpha_1^p\right)}$$

i.e., the planner's allocation α_1^p is recovered. Then we can solve b^A as $b^A = y_2 - \frac{y_1-y_2}{\frac{\left(1+2\alpha_1^p\right)\times\left(2-\alpha_1^p\right)\times\left(1-\alpha_2^p\right)}{\left(1+2\alpha_2^p\right)\times\left(2-\alpha_2^p\right)\times\left(1-\alpha_1^p\right)}-1}$. We can substitute y_1 for $y_1 = \frac{\alpha_1^p}{1-\alpha_1^p}y_2$. So that we could obtain $b^A = y_2 \times \left(1 - \frac{\frac{2\alpha_1^p-1}{1-\alpha_1^p}}{\frac{\left(1+2\alpha_1^p\right)\times\left(2-\alpha_1^p\right)\times\left(1-\alpha_2^p\right)}{\left(1+2\alpha_2^p\right)\times\left(2-\alpha_2^p\right)\times\left(1-\alpha_1^p\right)}-1}\right)$ where we express b^A as a function of α_1^p. It could be verified that this function of α_1^p is increasing in α_1^p. Besides, we know from $y_1 = \frac{\alpha_1^p}{1-\alpha_1^p}y_2$ that α_1^p is increasing in y_1, so that b^A is increasing in y_1. **Q.E.D.**

Proposition 4. *There exists a level of $\underline{w}^c \in (\hat{w}_2, y_2)$ such that the misallocation of surplus can be completely restored.*

Proof. We know that α_1^p is determined by $\frac{y_1}{y_2} = \frac{\alpha_1^p}{1-\alpha_1^p}$. According to $U^p = \left(1-\alpha_1^p\right)y_1 = \left(1-\alpha_2^p\right)y_2$, we are able to solve U^p as function of y_1 and y_2: $U^p = \frac{y_1 y_2}{y_1+y_2}$. We want to prove that under minimum wage policy, there is a level of minimum wage $\underline{w}^c < y_2$ such that the workers' expected surplus can be restored completely to the socially optimal level.

We now proceed to find the workers' expected surplus when the minimum wage policy is effective such that firm 2, which offers lower wage than firm 1, has to pay the workers the minimum wage $\underline{w}$. Then the firm 1 solves the following program:

$$\begin{array}{rl} \max_{w_1} & \left(1-(1-\alpha_1)^2\right)\times(y_1-w_1) \\ \text{s.to} & \left(1-\frac{\alpha_1}{2}\right)\times w_1=\left(1-\frac{\alpha_2}{2}\right)\times \underline{w} \end{array}$$

Since the constraint defines a one-one positive correlation between w_1 and α_1, we can also maximize with respect to α_1, which turns out to be more convenient. We could obtain as the first order condition $\frac{1-\hat{\alpha}_1^w}{1+2\hat{\alpha}_1^w}=\frac{\underline{w}}{y_1}\times\frac{1}{2}$, where $\hat{\alpha}_1^w$ denotes the equilibrium value of α_1 when minimum wage policy touches firm 2. We can solve the value of $\hat{\alpha}_1^w$, and can calculate the workers' expected utility as $U(\underline{w})=\left(1-\frac{\hat{\alpha}_2}{2}\right)\times\underline{w}=\frac{1}{4}\times\frac{4y_1+\underline{w}}{y_1+\underline{w}}\times\underline{w}$. This is an increasing function of $\underline{w}$.

Now it suffices to prove that for some feasible values of $\underline{w}$ with $\underline{w}<y_2$, it is possible that U^p is smaller than $U(\underline{w})$. For this sake, we can define a variable δ such that $\underline{w}=\delta y_2$ where $\delta\in(0,1)$. Then on one hand, we have $U^p=\frac{y_1y_2}{y_1+y_2}$, on the other hand, we have $U(\underline{w})=\frac{1}{4}\times\frac{4y_1+\delta y_2}{y_1+\delta y_2}\times\delta y_2=\frac{y_1\times\delta y_2+\frac{(\delta y_2)^2}{4}}{y_1+\delta y_2}$, which is greater than U^p when $\delta=1$. By continuity of δ, we could conclude that there exists a $\underline{w}^c$ such that workers' expected surplus can be restored. Q.E.D.

Corollary. *Introducing a proper level of unemployment benefit, together with minimum wage, can correct both the misallocation of labor and misallocation of surplus.*

Proof. When both unemployment benefit and minimum wage are introduced and the minimum wage is effective, expected payoff of workers from applying to firm 1 is $\left(1-\frac{\alpha_1}{2}\right)w_1+\frac{\alpha_1}{2}b$, and at firm 2 it is $\left(1-\frac{\alpha_2}{2}\right)\underline{w}+\frac{\alpha_2}{2}b$. Firm 1 solves the following program:

$$\begin{array}{rl} \max_{w_1} & \left(1-(1-\alpha_1)^2\right)\times(y_1-w_1) \\ \text{s.to} & \left(1-\frac{\alpha_1}{2}\right)\times(w_1-b)=\left(1-\frac{\alpha_2}{2}\right)\times(\underline{w}-b) \end{array}$$

Maximizing with respect to α_1, we obtain as the first order condition $\frac{1-\tilde{\alpha}_1}{1+2\tilde{\alpha}_1}=\frac{\underline{w}-b}{y_1-b}\times\frac{1}{2}$, where $\tilde{\alpha}_1$ denotes the equilibrium value of α_1. This defines $\tilde{\alpha}_1$ as a function of the policy instruments $\underline{w}$ and b. We impose $\tilde{\alpha}_1=\alpha_1^p$ so that the misallocation of labor is corrected, then we obtain a first relationship between $\underline{w}$ and b.

We now want to find the second relation, and it is done by imposing the expected payoff to be equivalent to its socially optimal level: the misallocation of surplus is corrected. We calculate the workers' expected payoff: $U(\underline{w})=\left(1-\frac{\tilde{\alpha}_2}{2}\right)\times\underline{w}+\frac{\tilde{\alpha}_2}{2}b=\left(1-\frac{\tilde{\alpha}_2}{2}\right)\times(\underline{w}-b)+b$. We let it be equal to the socially optimal payoff $U^p=\left(1-\alpha_1^p\right)y_1$. So that we have $\left(1-\frac{\tilde{\alpha}_2}{2}\right)\times(\underline{w}-b)+b=\left(1-\alpha_1^p\right)y_1$.

Rearranging, it suffices that the system of equations where $\underline{w}$ and b serve as variables admits a solution:

$$\begin{cases} \left(\frac{1+2\alpha_1^p}{1-4\alpha_1^p}\right)\underline{w}+b = \frac{2\times\left(1-\alpha_1^p\right)}{1-4\alpha_1^p}y_1 \\ \left(\frac{1+\alpha_1^p}{1-\alpha_1^p}\right)\underline{w}+b = 2y_1 \end{cases}$$

We are able to obtain $b = \frac{\frac{2}{1+\alpha_1^p}-\frac{2}{1+2\alpha_1^p}}{\left(\frac{1}{1+\alpha_1^p}+\frac{1}{1-\alpha_1^p}\right)-\left(\frac{2}{1+2\alpha_1^p}\right)}y_1$ and $\underline{w} = \frac{2\times\left(1-\frac{1-\alpha_1^p}{1-4\alpha_1^p}\right)}{\left(\frac{1+\alpha_1^p}{1-\alpha_1^p}\right)-\left(\frac{1+2\alpha_1^p}{1-4\alpha_1^p}\right)}y_1$. It could be checked that $\underline{w} > b > 0$. Q.E.D.

Proposition 5. *In system (B), there exists no unemployment benefit level b^B such that the efficiency is restored in the sense of GKV(2011).*

Proof. Under unemployment benefit system B, firm 1 pays a level of unemployment benefit $(\alpha_1)^2 b$, and firm 2 pays a level of unemployment benefit $(\alpha_2)^2 b$, where $(\alpha_1)^2$ and $(\alpha_2)^2$ are respectively the probability of paying b to the workers who applied but did not get the job. This is quite intuitive: in expectation, a firm has to pay this b only when both workers apply to it but one among them will not be selected, an event which takes place with probability α^2. Hence the problems for the firms are written as

$$\begin{array}{ll} \max_{w_1} & \left(1-(1-\alpha_1)^2\right)\times(y_1-w_1)-(\alpha_1)^2 b \\ \text{s.to} & \frac{2-\alpha_1}{2}\times(w_1-b) = \frac{2-\alpha_2}{2}\times(w_2-b) \end{array}$$

and

$$\begin{array}{ll} \max_{w_2} & \left(1-(1-\alpha_2)^2\right)\times(y_2-w_2)-(\alpha_2)^2 b \\ \text{s.to} & \frac{2-\alpha_1}{2}\times(w_1-b) = \frac{2-\alpha_2}{2}\times(w_2-b) \end{array}$$

The first order conditions now become

$$2(1-\hat{\alpha}_1)\times(y_1-b)-(1+2\hat{\alpha}_1)\times(w_2-b)-(2\hat{\alpha}_1)b = 0$$

and

$$2(1-\hat{\alpha}_2)\times(y_2-b)-(1+2\hat{\alpha}_2)\times(w_1-b)-(2\hat{\alpha}_2)b = 0$$

Combining these two, we obtain

$$\frac{y_1 - \frac{1}{1-\hat{\alpha}_1^b} b}{y_2 - \frac{1}{1-\hat{\alpha}_2^b} b} = \frac{\left(1+2\hat{\alpha}_1^b\right) \times \left(2-\hat{\alpha}_1^b\right) \times \left(1-\hat{\alpha}_2^b\right)}{\left(1+2\hat{\alpha}_2^b\right) \times \left(2-\hat{\alpha}_2^b\right) \times \left(1-\hat{\alpha}_1^b\right)}$$

It is more useful to write this condition as

$$\frac{\left(1-\hat{\alpha}_1^b\right) y_1 - b}{\left(1-\hat{\alpha}_2^b\right) y_2 - b} = \frac{\left(1+2\hat{\alpha}_1^b\right) \times \left(2-\hat{\alpha}_1^b\right)}{\left(1+2\hat{\alpha}_2^b\right) \times \left(2-\hat{\alpha}_2^b\right)}$$

Suppose there is a b^B such that $\hat{\alpha}_1^b$ could be restored to the socially optimal level α_1^p, then the above equality becomes

$$\frac{\left(1-\alpha_1^p\right) y_1 - b^B}{\left(1-\alpha_2^p\right) y_2 - b^B} = \frac{\left(1+2\alpha_1^p\right) \times \left(2-\alpha_1^p\right)}{\left(1+2\alpha_2^p\right) \times \left(2-\alpha_2^p\right)}$$

Since we have $\left(1-\alpha_1^p\right) y_1 = \left(1-\alpha_2^p\right) y_2$. It requires that $\frac{\left(1+2\alpha_1^p\right) \times \left(2-\alpha_1^p\right)}{\left(1+2\alpha_2^p\right) \times \left(2-\alpha_2^p\right)} = 1$ which is impossible because $\frac{\left(1+2\alpha_1^p\right) \times \left(2-\alpha_1^p\right)}{\left(1+2\alpha_2^p\right) \times \left(2-\alpha_2^p\right)}$ is always greater than 1 when $\alpha_1^p > \frac{1}{2} > \alpha_2^p$. Q.E.D.

Claim. With fixed sharing rule β^{NB} such that workers' expected payoff is equal to U^p,

(1) the high (low) productivity firm offers higher (lower) wage compared to the case where wages are posted.

(2) both the high and low productivity firm earn lower profit compared to the case where wages are posted.

Proof of (1). We start to compare $\beta^{\mathrm{NB}} y_1$ and w_1^p. Recall that $w_1^p = \frac{2}{2-\alpha_1^p}\left(1-\alpha_1^p\right) y_1$, and $\beta^{\mathrm{NB}} y_1 = \frac{2}{2-\alpha_1^{\mathrm{NB}}} \times \left(1-\alpha_1^p\right) y_1$. Since $\alpha_1^{\mathrm{NB}} > \alpha_1^p$, we have $w_1^p < \beta^{\mathrm{NB}} y_1$. We now compare $\beta^{\mathrm{NB}} y_2$ and w_2^p. Recall that $w_2^p = \frac{2}{2-\alpha_2^p}\left(1-\alpha_2^p\right) y_2$, and $\beta^{\mathrm{NB}} y_2 = \frac{2}{2-\alpha_2^{\mathrm{NB}}} \times \left(1-\alpha_2^p\right) y_2$. Since $\alpha_2^{\mathrm{NB}} < \alpha_2^p$, we have $w_2^p > \beta^{\mathrm{NB}} y_2$.

Proof of (2). We prove the result for firm 1, and the proof for firm 2 will be analogous. We know that w_1^p and α_1^p solve the following program for firm 1:

$$\begin{array}{ll} \max & \left(1-\left(1-\alpha_1\right)^2\right)\left(y_1-w_1\right) \\ \text{s.to} & \frac{2-\alpha_1}{2} \times w_1 = U^p \end{array}$$

Define $w_1^{\text{NB}} = \beta^{\text{NB}} y_1$. We also know that α_1^{NB} and w_1^{NB} also satisfies the constraint, however α_1^{NB} is different from α_1^p, and w_1^p is different from α_1^p. By this, we are able to conclude that

$$\begin{aligned} & \left(1-\left(1-\alpha_1^p\right)^2\right)\left(y_1-w_1^p\right) \\ > & \left(1-\left(1-\alpha_1^{\text{NB}}\right)^2\right)\left(y_1-w_1^{\text{NB}}\right) \end{aligned}$$

Q.E.D.

References

[1] Acemoglu, D. and Shimer, R. (1999). Efficient unemployment insurance. *Journal of Political Economy*, 107(5):893–928.

[2] Acemoglu, D. and Shimer, R. (1999b). Holdups and efficiency with search frictions. *International Economic Review*, 40(4):827–849.

[3] Albrecht, J., Tan, S., Gautier, P., and Vroman, S. (2004). Matching with multiple applications revisited. *Economics Letters*, 84(3):311–314.

[4] Altonji, J. G. and Williams, N. (1998). The effects of labor market experience, job seniority and job mobility on wage growth. *Research in Labor Economics*, 17.

[5] Bagger, J., Fontaine, F., Postel-Vinay, F., and Robin, J.-M. (2014). Tenure, experience, human capital, and wages: A tractable equilibrium search model of wage dynamics. *The American Economic Review*, 104(6):1551–1596.

[6] Bi, S. and Langot, F. (2015). Dynamic contracts in search equilibrium under asymmetric information.

[7] Brenzel, H., Gartner, H., and Schnabel, C. (2014). Wage bargaining or wage posting? evidence from the employers' side. *Labour Economics*, 29:41–48.

[8] Burdett, K., Shi, S., and Wright, R. (2001). Pricing and matching with frictions. *Journal of Political Economy*, 109(5):1060–1085.

[9] Cao, M. and Shi, S. (2000). Coordination, matching, and wages. *Canadian Journal of Economics*, pages 1009–1033.

[10] Chéron, A., Hairault, J.-O., and Langot, F. (2011). Age-dependent employment protection*. *The Economic Journal*, 121(557):1477–1504.

[11] Chéron, A., Hairault, J.-O., and Langot, F. (2013). Life-cycle equilibrium unemployment. *Journal of Labor Economics*, 31(4):843–882.

[12] Coile, C. C. and Levine, P. B. (2006). Bulls, bears, and retirement behavior. *Industrial & Labor Relations Review*, 59(3):408–429.

[13] Coile, C. C. and Levine, P. B. (2007). Labor market shocks and retirement: Do government programs matter? *Journal of Public Economics*, 91(10):1902–1919.

[14] Coile, C. C. and Levine, P. B. (2011a). The market crash and mass layoffs: How the current economic crisis may affect retirement. *The BE Journal of Economic Analysis & Policy*, 11(1).

[15] Coile, C. C. and Levine, P. B. (2011b). Recessions, retirement, and social security. *The American Economic Review*, pages 23–28.

[16] Cutler, D., Meara, E., and Richards-Shubik, S. (2011). Healthy life expectancy: Estimates and implications for retirement age policy. Technical report, National Bureau of Economic Research.

[17] Delacroix, A. and Shi, S. (2013). Pricing and signaling with frictions. *Journal of Economic Theory*, 148(4):1301–1332.

[18] Fang, H. and Moro, A. (2010). Theories of statistical discrimination and affirmative action: A survey. Technical report, National Bureau of Economic Research.

[19] Farber, H. S. (1994). The analysis of interfirm worker mobility. *Journal of Labor Economics*, pages 554–593.

[20] Galenianos, M. and Kircher, P. (2009). Directed search with multiple job applications. *Journal of Economic Theory*, 144(2):445–471.

[21] Galenianos, M. and Kircher, P. (2012). On the game-theoretic foundations of competitive search equilibrium*. *International economic review*, 53(1):1–21.

[22] Galenianos, M., Kircher, P., and Virág, G. (2011). Market power and efficiency in a search model*. *International economic review*, 52(1):85–103.

[23] Geromichalos, A. (2012). Directed search and optimal production. *Journal of Economic Theory*, 147(6):2303–2331.

[24] Geromichalos, A. (2014). Unemployment insurance and optimal taxation in search models of the labor market. Technical report, Working Papers, University of California, Department of Economics.

[25] Grout, P. A. (1984). Investment and wages in the absence of binding contracts: A nash bargaining approach. *Econometrica*.

[26] Guerrieri, V. (2008). Inefficient unemployment dynamics under asymmetric information. *Journal of Political Economy*, 116(4):667–708.

[27] GUERRIERI, V., SHIMER, R., and WRIGHT, R. (2010). Adverse selection in competitive search equilibrium. *Econometrica*, 78(6):1823–1862.

[28] Hairault, Jean-Olivier, F. L. and Zylberberg., A. (2015). Equilibrium unemployment and retirement. *European Economic Review*.

[29] Hairault, J.-O., Langot, F., Ménard, S., and Sopraseuth, T. (2012). Optimal unemployment insurance for older workers. *Journal of Public Economics*, 96(5):509–519.

[30] Hairault, J.-O., Sopraseuth, T., and Langot, F. (2010). Distance to retirement and older workers 'employment: The case for delaying the retirement age. *Journal of the European Economic association*, 8(5):1034–1076.

[31] Hall, R. E. and Krueger, A. B. (2012). Evidence on the incidence of wage posting, wage bargaining, and on-the-job search. *American Economic Journal: Macroeconomics*, pages 56–67.

[32] Heckman, J. J. (1998). Detecting discrimination. *The Journal of Economic Perspectives*, pages 101–116.

[33] Hopenhayn, H. A. and Nicolini, J. P. (1997). Optimal unemployment insurance. *Journal of political economy*, 105(2):412–438.

[34] Jacquet, N. L. and Tan, S. (2012). Wage-vacancy contracts and coordination frictions. *Journal of Economic Theory*, 147(3):1064–1104.

[35] Laffont, J.-J. and Martimort, D. (2001). The theory of incentives: The principal-agent model.

[36] Lang, K. and Lehmann, J.-Y. K. (2012). Racial discrimination in the labor market: Theory and empirics. *Journal of Economic Literature*, 50(4):1–48.

[37] Lang, K. and Manove, M. (2003). Wage announcements with a continuum of worker types. *Annales d'Economie et de Statistique*, pages 223–244.

[38] Lang, K. and Manove, M. (2011). Education and labor market discrimination. *American Economic Review*, 101(4):1467–96.

[39] Lang, K., Manove, M., and Dickens, W. T. (2005). Racial discrimination in labor markets with posted wage offers. *American Economic Review*, pages 1327–1340.

[40] Lester, B. (2010). Directed search with multi-vacancy firms. *Journal of Economic Theory*, 145(6):2108–2132.

[41] Li, F. and Tian, C. (2013). Directed search and job rotation. *Journal of Economic Theory*, 148(3):1268–1281.

[42] Low, H. and Pistaferri, L. (2010). Disability risk, disability insurance and life cycle behavior. Technical report, National Bureau of Economic Research.

[43] Menzio, G., Telyukova, I. A., and Visschers, L. (2015). Directed search over the life cycle. *Review of Economic Dynamics*.

[44] Merlino, L. P. (2012). Discrimination, technology and unemployment. *Labour economics*, 19(4):557–567.

[45] Michelacci, C. and Ruffo, H. (2015). Optimal life cycle unemployment insurance. *American Economic Review*, 105(2):816–859.

[46] Moen, E. R. (1997). Competitive search equilibrium. *Journal of political Economy*, 105(2):385–411.

[47] Mortensen, D. T. (2011). Autobiography of dale t. mortensen.

[48] Mortensen, D. T. and Pissarides, C. A. (1994). Job creation and job destruction in the theory of unemployment. *The review of economic studies*, 61(3):397–415.

[49] Neal, D. A. and Johnson, W. R. (1996). The role of premarket factors in black-white wage differences. *The Journal of Political Economy*, 104(5):869–895.

[50] O'Neill, J. (1990). The role of human capital in earnings differences between black and white men. *The Journal of Economic Perspectives*, pages 25–45.

[51] Pendakur, K. and Woodcock, S. (2010). Glass ceilings or glass doors? wage disparity within and between firms. *Journal of Business & Economic Statistics*, 28(1):181–189.

[52] Peters, M. (2010). Noncontractible heterogeneity in directed search. *Econometrica*, 78(4):1173–1200.

[53] Petrongolo, B. and Pissarides, C. A. (2001). Looking into the black box: A survey of the matching function. *Journal of Economic literature*, pages 390–431.

[54] Phelps, E. S. (1970). Microeconomic foundations of employment and inflation theory.

[55] Ritter, J. A. and Taylor, L. J. (2011). Racial disparity in unemployment. *The Review of Economics and Statistics*, 93(1):30–42.

[56] Rosén, Å. (1997). An equilibrium search-matching model of discrimination. *European Economic Review*, 41(8):1589–1613.

[57] Rothschild, M. and Stiglitz, J. (1976). Equilibrium in competitive insurance markets: An essay on the economics of imperfect information. *The Quarterly Journal of Economics*, 90(4):629–649.

[58] Salop, J. and Salop, S. (1976). Self-selection and turnover in the labor market. *The Quarterly Journal of Economics*, pages 619–627.

[59] Shavell, S. and Weiss, L. (1979). The optimal payment of unemployment insurance benefits over time. *The Journal of Political Economy*, pages 1347–1362.

[60] Shi, S. (2001). Frictional assignment. i. efficiency. *Journal of Economic Theory*, 98(2):232–260.

[61] Shi, S. (2002a). A directed search model of inequality with heterogeneous skills and skill-biased technology. *The Review of Economic Studies*, 69(2):467–491.

[62] Shi, S. (2002b). Product market and the size–wage differential*. *International Economic Review*, 43(1):21–54.

[63] Shi, S. (2006). Wage differentials, discrimination and efficiency. *European Economic Review*, 50(4):849–875.

[64] Shimer, R. (2005). The assignment of workers to jobs in an economy with coordination frictions. *Journal of Political Economy*, 113(5):996–1025.

[65] Shimer, R. (2007). Mismatch. *The American Economic Review*, 97(4):1074.

[66] Tan, S. (2012). Directed search and firm size*. *International Economic Review*, 53(1):95–113.

[67] Virág, G. (2011). High profit equilibria in directed search models. *Games and Economic Behavior*, 71(1):224–234.