Ranking and Search Effort in Matching

Joonbae Lee[*], Hanna Wang[†]

January 25, 2018

Abstract

This paper studies the relationship between search effort and worker's rank in a random search environment with fixed wages. We build a matching model in which firms' common preferences over a continuum of heterogeneous workers affect the number of applications each type sends out in the hiring process. We show that in equilibrium the relationship between rank and number of applications must not be monotonous. In fact, we prove that it is hump-shaped for sufficiently high vacancy-to-worker ratios, that is highly-ranked and lowly-ranked workers send out fewer applications than workers of mid-range rank. This arises due to two opposing forces driving the incentive of applicants. Increasing the number of applications acts as insurance against unemployment, but is less effective when the probability of success for each application is low. This mechanism exacerbates the negative employment outcomes of low rank workers. We discuss comparative statics with regards to the size of the vacancy pool and application cost, and show that the solution of the social planner has the number of applications be monotonously decreasing in rank.

Keywords Simultaneous Search, Search Friction, Occupancy Problem

1 Introduction

Mismatch in labor markets with simultaneous applications is a widely-studied phenomenon. Starting from Mortensen (1985), Diamond (1982) and Pissarides (1990), macroeconomic search models typically represent matching friction using matching functions. The functions map market tightness into a number of matches, that is always less than the number

[*]joonbae@sas.upenn.edu
[†]hannaw@sas.upenn.edu

of vacancies or unemployed workers. The idea behind it is simple: Without coordination, many workers apply to the same firms while other firms do not receive any applications, resulting in unemployment and unfilled vacancies.

The extent of this miscoordination affects the success probability of worker applications and worker's incentive to search for jobs. Dating back to Stigler (1962) the search and matching literature has studied how endogenous search effort affect labor market outcomes, namely through sampling for better wage offers and increasing the probability of finding any job. Search models typically model search effort as a continuous variable that is multiplied by the arrival rate (e.g. Pissarides (2000)) while matching models incorporate it as the choice of a discrete number of applications. While most paper consider settings with homogeneous workers, it is natural to imagine that workers - even within the same applicant pool - can be different from one another and hence face different conditions in the application process. This paper provides an analysis of number of applications sent out by workers of different rank and their probability of securing a job from a pool of identical jobs.

Heterogeneity on the searcher side has commonly only been studied under the assumption of a fixed level of search effort, e.g. in matching models with directed applications such as in Peters (2010), Shimer (2005) and Shi (2002), or in search models with directed search and/or sorting such as Menzio and Shi (2011) and Burdett and Coles (1997). A few search papers develop models with worker differentiation and endogenous search, e.g. Lentz (2010), but focus on wage increases through on-the-job search. As Shimer (2004) pointed out, traditional search models generate the result that equilibrium search effort has to increase in the baseline arrival rate. This would suggest that higher-rank workers search more given that they are favored over other workers. In contrast, we show in our model that the relationship between rank and search effort is hump-shaped, so for workers above a certain threshold rank search effort decreases with rank, while the reverse is true for workers below the threshold.

At first glance the effect of rank on the number of applications is not obvious. In particular, there are two opposing incentives if we compare a low-rank to a high-rank worker. On the one hand, because success probability per application decreases with rank (as can be immediately shown in our model) and applications are costly, the low-rank worker might be discouraged from applying as often as the high-rank worker. This is what a search model would predict. On the other hand, he might want to send out additional applications as a form of insurance, because previous applications are more likely to fail. For instance, the most preferred worker would only ever want to send out one application, since it will land him a job with certainty. Our main result implies that above a threshold rank, lower rank workers can gain on, if not overtake, higher rank workers in terms of employment probability, by counteracting lower success rate per application with sending more applications. However, below the threshold, employment chance rapidly decreases with rank due to workers both facing an even lower per-application success rate and sending out fewer applications than higher-ranked competitors.

We model worker heterogeneity as a characteristic over which firms have common preferences but cannot contract. Hence, firms cannot discriminate between workers by posting wage schedules based on types but only by selecting a preferred applicant in the hiring process. This is assumption mirrors Peters (2010) and simplifies the comparison of incentives across types. We further assume that the market has a fixed wage, so that workers cannot distinguish among vacancies. This intensifies competition across worker types since workers cannot evade one another by directing their applications. Our setting can be interpreted in different ways. One possibility is that the value of the worker characteristic is difficult to quantify, e.g. soft skills or reference letters, or it is not socially acceptable or legal to base wage contracts on it, e.g. race, sex, motherhood status or physical attractiveness. Another case where our setting applies are labor markets with fixed wages, for instance the market for minimum wages jobs or markets with standardized uniform starting salaries for workers with given certain qualifications, e.g. college graduates. It is conceivable that in reality, even in very specialized or small markets such characteristics by which applicants can be ranked, exist to some extent.

Modelling multiple applications per worker can be a challenging task due to the fact that firms and workers actions are interdependent on one another and characterizing employment probabilities involves cumbersome combinatorics. To see why note that a firm can only hire a particular worker if he does not decide to work at another firm and similarly, a worker can only work at a particular firm if that firm does not hire a different worker. To limit the rounds of back-and-forth offers and acceptances, the majority of papers restrict firms to make a single offer, e.g. Shimer (2004), Albrecht, Gautier and Vroman (2003) and Galenianos and Kircher (2009). The equilibrium notion of stable matching was first introduced in Kircher (2009) and effectively allows for workers and firms to communicate back and forth until all possible matches are realized. Our paper joins a handful of papers (Gautier and Holzner (2013, 2016), Wolthoff (2009)) that have adapted this equilibrium notion. In contrast to our paper, the papers listed above are motivated by efficiency concerns in the presence of wage dispersion, for instance Kircher shows that equilibria in stable matching models are constrained efficient where as those in one-shot firm offer models are not.

A stable matching equilibrium rules out strictly profitable deviations for any worker and firm. In our context, on the firm side, firms are cannot have unemployed applicants while having an open vacancy or hiring a worker of lower rank than one of the unemployed applicants. On the worker side unemployment is only possible when all firms the worker applied to hire higher-ranked workers (we let the payoffs to be such that the worker prefers employment over unemployment). Hence each worker will be employed if at least one of the firms he applied to is not able to hire a higher-ranked worker. Given this and a continuum worker type setting we can obtain well-defined expressions for each worker's employment probability as a function of the deterministic mass of employed higher-rank workers. Furthermore, we derive a differential equation that characterize how the latter evolves with worker rank.

The paper proceeds as follows. We illustrate the basic idea using a simple example in Section 2. Section 3 develops the full theoretical model and presents the main result. We then study how market tightness and application costs affects application behaviour and employment outcomes in Sections 4 and 5, and discuss efficiency and the social planner solution in section 6.

2 A Simple Discrete Example

Consider a labour market with M vacancies, each at a different firm, and three workers, $\{A, B, C\}$, ranked first, second, and third respectively by all firms. The workers decide how many applications to send to the firms, which can be thought of as a random sampling out of the M jobs. A worker can not send multiple applications to a same job, so he samples jobs without replacement. Therefore, for a sample size of m, a worker has $\binom{M}{m}$ ways of sampling firms. A worker occupies at most one job among the sampled ones.

Workers do not know where the others send the applications to, and there is a possibility that a competitor is chosen over him. This coordination failure is resolved by the rank order of the workers, and the market clears in following manner: After the applications have been sent, the highest rank worker chooses, among the jobs he sampled, which job to accept. After that, the second highest rank worker chooses among his sampled jobs, however he cannot choose a job that has already been chosen by the highest ranked worker. Hence, if he only applied once, and that firm happens to employ the highest rank worker, he will remain unemployed. If he sent out at least two applications he will get hired for sure. Lastly, the lowest rank worker chooses where to go; again if all of the jobs he sampled are already taken by the first two workers, the worker is left unemployed. If there are at least one vacancy that is unfilled, the worker succeeds in finding a job. In fact, if the worker has sent out many applications, the worker might enjoy a chance of having to choose among multiple positions.

The benefit to sending out more applications is obvious: by sampling more jobs, the worker decreases the probability that his applications overlap with higher rank workers, and thereby increases the probability of finding a vacant job. Furthermore, since the market clears top-down, it suffices for a worker to calculate the probability that his application coincides with successful applications of higher rank workers. We also note that the process does not maximize the expected number of matches.

Note that worker A will send out at most one application, his application will always be successful.

For $M = 2$, the probability of each worker getting a job, as a function of applications sent out by the workers, is as given in the following table, fixing worker A's number applications to one.

(B,C)	1	2
1	$(\frac{1}{2}, \frac{1}{2}\frac{1}{2} + \frac{1}{2} \cdot 0)$	$(\frac{1}{2}, \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot 0)$
2	$(1, 0)$	$(1, 0)$

From the table, we read off the marginal benefit of additional application for each type:

MB	$1 - 0$	$2 - 1$
A	1	0
B	$\frac{1}{2}$	$\frac{1}{2}$
C (1)	$\frac{1}{4}$	$\frac{1}{4}$
C (2)	0	0

The third row shows worker C's marginal benefit when the worker B sends out one application; while the fourth row is the counterpart when the worker B sends out two applications.

We see the marginal benefit of application for worker C is a function of worker B's application. Given that worker A gets a job for sure, when worker B sends out one application he might not become employed, leaving one potential job for worker C. If worker B sends out two applications, worker C has no chance of receiving a job with any number of applications.

We denote the flat cost of applications as c and specify the worker's expected utility as the probability of getting a job multiplied by the wage minus the cost of application. The workers send additional application as long as the marginal benefit of the application exceeds the cost. In this simple example, if the wage is one for all jobs, we can see that for application cost less than $\frac{1}{2}$ worker A applies once, worker B applies twice, while worker C does not apply at all. This highlights the intuition behind our main result that the number of applications is hump-shaped in rank. Workers with a mid-range rank might apply more to insure against losing against higher rank competitors while low rank workers might apply less because too many jobs are taken by higher rank workers.

Now, to preview our comparative statics results, consider the following changes. First, the market becomes less tight for workers; there are now $M = 3$ jobs. The table for probability is now given:

(B,C)	1	2	3
1	$(\frac{2}{3}, \frac{2}{3}\frac{1}{3} + \frac{1}{3}\frac{2}{3})$	$(\frac{2}{3}, \frac{2}{3}\frac{2}{3} + \frac{1}{3} \cdot 1)$	$(\frac{2}{3}, 1)$
2	$(1, \frac{1}{3})$	$(1, \frac{1}{3} \cdot 0 + \frac{2}{3})$	$(1, 1)$
3	$(1, \frac{1}{3})$	$(1, \frac{1}{3} \cdot 0 + \frac{2}{3})$	$(1, 1)$

Comparing with the corresponding components in the table for $M = 2$, we observe that the payoffs uniformly increased for both workers, which is intuitive. The marginal benefits are now:

MB	$1-0$	$2-1$	$3-2$
A	1	0	0
B	$\frac{2}{3}$	$\frac{1}{3}$	0
C (1)	$\frac{4}{9}$	$\frac{3}{9}$	$\frac{2}{9}$
C (2)	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

We see that the effect of increasing M on marginal benefit of applications is ambiguous. For instance, worker B's marginal benefit for the first application increased, while the second application has become less valuable.

Hence, we see that for some cost parameter c, the effect on the number of applications of increasing M may go in different directions. For instance, for c between $\frac{1}{3}$ and $\frac{4}{9}$, worker B will decrease his number of applications from two to one, but worker C increases his application from zero to one. Intuitively, in response to a more favorable market condition, workers ranked in the mid-range decrease their search effort, since their insurance motive has diminished, while low rank workers increase their search effort, since more they are not as discouraged.

We also observe that decreasing the cost of application might hurt worker C in the equilibrium. For the above range of c the expected benefits are $\frac{2}{3} - c < \frac{1}{3}$ for worker C, and $\frac{2}{3}\frac{1}{3} + \frac{1}{3}\frac{2}{3} - c = \frac{4}{9} - c < \frac{1}{9}$ for worker B. However, if c is slightly below $\frac{1}{3}$, worker B will now send out two applications, and worker C (reading from the fourth row of the table) will send three applications. The utility is $1 - 2c > \frac{1}{3}$ for worker B, and $1 - 3c > 0$ for worker C. For c close enough to $\frac{1}{3}$, we observe that worker B gained from the change in cost, while worker C is worse off.

3 Model

Setting We consider a bipartite matching problem of two sets of continuum of agents: workers $\mathcal{W}$, and firms $\mathcal{F}$. Let $\mathcal{W}$ be a unit mass of workers, indexed by $i \in [0, 1]$. Each worker is endowed with type $x \in [0, 1]$, and we denote workers' type distribution by F with support $[0, 1]$ and a pdf f. Define $X : [0, 1] \to [0, 1]$ which maps each worker to her type. Then F is induced by X, and Lebesgue measure λ over $\mathcal{W}$: $F(x) := \lambda(\{i : X(i) \leq x\})$. Let $\mathcal{F}$ be a positive mass $M(> 0)$ of firms indexed by $j \in [0, M]$. Firms are identical and each firm has one vacancy.

Firms match with workers through application and hiring process. A matching game in this market consists of two-stages. In the first stage, called "application process," each worker simultaneously chooses an integer $k \in \mathbb{Z}_+$, which represents the number of applications to send out to firms. This the only strategic component in this matching game. If a worker chose to send k applications, then the worker makes k i.i.d. uniform draws from the firm pool ranging over $[0, M]$. We think of this as workers choosing the number of 'links' to form with the set $\mathcal{F}$, while the actual identity of the firms to link with is random.

In the second stage, matches between firms and workers, who are connected through a realized application, are formed. We describe the matching process in the second stage in detail below. First, we outline the payoffs of the game. Since firms are identical, we normalize the worker's gain from matching with a firm one.

If a worker remains unmatched, his payoff is 0. Each application a worker sends incurs cost c. A firm's payoff is equal to the matched worker's type, or $-\varepsilon(< 0)$ if it remains unmatched. A firm strictly prefers to be matched with any worker than remaining unmatched.

Matching Process The application and hiring process begins with all workers simultaneously choosing how many applications to send out. Each worker sends her applications to randomly selected firms. It follows that some firms may receive multiple applications, while others receive none.

We define a *matching* as a function μ: $\mathcal{W} \cup \mathcal{F} \to \mathcal{W} \cup \mathcal{F}$ such that

1. $\mu(i) = j \in \mathcal{F}$ if and only if $\mu(j) = i \in \mathcal{W}$,
2. If $\mu(i) \notin \mathcal{F}$, then $\mu(i) = i$,
3. If $\mu(j) \notin \mathcal{W}$, then $\mu(j) = j$.

For each $i \in \mathcal{W}$, and corresponding $k(i)$ applications i sent out in the first stage, define the set of firms receiving an application from i as $B(i) = \{j_1, j_2, \ldots, j_{k(i)}\}$. A matching μ in the second stage is *feasible* if $\mu(i) \in \{i\} \cup B(i)$ for all i. We require the second stage matching μ to be feasible, and *stable* in the following sense.

Definition 1. A feasible matching μ is stable if there is no pair (i, j) with $j \in B(i)$ such that $\mu(i) = i$ and either (i) $\mu(j) = j$, or (ii) $\mu(j) \neq j$ and $X(i) > X(\mu(j))$.

This is the notion of no blocking pair for which both parties are strictly better off. For any blocking pair (i, j) with $j \in B(i)$, since workers are indifferent over all firms, it must be $\mu(i) = i$. Furthermore, j strictly prefers i to his current match $\mu(j)$, himself or a worker with type lower than i.

This characterizes all stable matchings resulting from a given realization of applications. Note that there might be multiple and we will use the following procedure to select one:

- The highest type worker i_1 with $X(i_1) = 1$ chooses at random $\mu(i_1) \in B(i_1)$ if $B(i_1)$ is non-empty.

- The next highest worker, i_2, chooses at random $\mu(i_2) \in B(i_2)\backslash\mu(i_1)$, if this set is non-empty.
- remaining workers choose in a similar fashion, with order descending in type.

Note that it does not select the matching which maximizes the number of matches, because workers randomize when choosing an offer and do not take into consideration how this affects lower rank workers. It can be seen that the clearing process is equivalent to randomly selecting one stable matching among all possible stable matchings given one realization of applications.

Matching Probability In this section, we examine the consequences on the employment probability of the aforementioned matching process. For illustration, consider an economy with a finite number of workers and m firms. For now assume each of the n workers send out only one application, then the number of their applications at a specific firm is a random variable that follows a binomial distribution, $X \sim bin(n-1, \frac{1}{m})$. As we increase n, m to infinity keeping the ratio constant, $\frac{n}{m} \to \lambda$, the well-known result in the limit approximation of binomial random variable tells us that the limit X follows Poisson with parameter λ.

In our model, for a type x worker there are $1-F(x)$ workers of higher type. Using the previous argument the number of applications of higher types follows a Poisson distribution with parameter $\frac{1-F(x)}{M}$. Since everyone sends out one application, the worker gets hired if and only if his application do not overlap with an application of a higher type. The probability corresponds to the zero realization of the Poisson random variable. The employment probability is $e^{-\frac{1-F(x)}{M}}$.

Assuming that the derivative of F, f, exists everywhere and is bounded, we note that the probability can be equally derived from the unique solution to the following differential equation

$$-A'(x) = f(x)\left(1 - \frac{A(x)}{M}\right)$$

with initial value $A(1) = 0$, according to the relation

$$1 - \frac{A(x)}{M} = e^{-\frac{1-F(x)}{M}}.$$

This differential equation has a very intuitive interpretation. Note that the expression on the right-hand-side is the probability that type x is hired ($1 - \frac{A(x)}{M} = e^{-\frac{1-F(x)}{M}}$) multiplied by the pdf evaluated at x, hence we state this as the change in A is equal to the expected mass of type x workers who succeeds in finding a job. This allows us to interpret $A(x)$ as the mass of workers up until type x who landed a job. In fact, the differential equation can be derived as a limit of a finite type model, as the types become infinitesimal. A detailed discussion is relegated to Appendix A.1.

Now we are left to check whether this intuition extends to multiple applications. According to our interpretation of A, the differential equation is modified to:

$$-A'(x) = f(x)\left(1 - \left(\frac{A(x)}{M}\right)^k\right)$$

given that $1 - \left(\frac{A(x)}{M}\right)^k$ is the probability that a x worker finds a job. Given the top-down clearing process, we argue that the probability of each worker can be expressed as a function of the mass of employed higher-type workers. We elaborate in Appendix A.2.

With multiple applications even if all of a worker's applications overlap with those of higher type workers, he needs not remain unemployed. This is because higher type competitors might accept other jobs, too, if they sent out more than one application. Due to random sampling of jobs, the probability that each of the application succeeds is independent.

Equilibrium The equilibrium of this game, given constant cost c, mass of firms M, and the distribution of worker types F, are functions (A, k) such that

- Given a nondecreasing $A : [0,1] \to [0, M]$, $k : [0,1] \to \mathbb{N} \cup \{0\}$ pointwise maximizes the utility of the worker. That is, for all x:

$$k(x) = \arg \max_{k \in \mathbb{N} \cup \{0\}} 1 - \left(\frac{A(x)}{M}\right)^k - ck$$

- Given k, A evolves according to the differential equation:

$$-A'(x) = f(x)(1 - \left(\frac{A(x)}{M}\right)^{k(x)}), \tag{1}$$

  with initial condition $A(1) = 0$.

Intuitively, the equilibrium definition incorporates that each worker's application behavior is optimal given the application behavior of all other workers. Since the market clears top

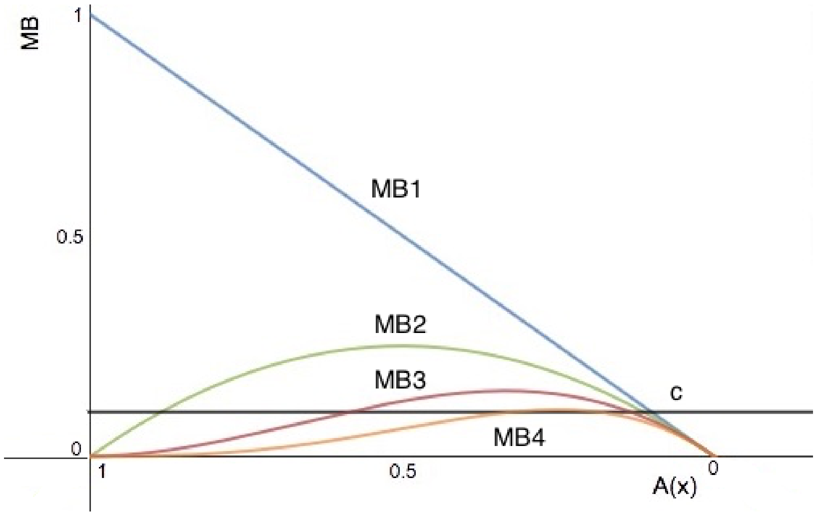

Figure 1: Marginal benefits of applications 1-4

down, the worker need only worry about sending applications to jobs that end up hiring higher rank workers. $A(x)$ is a sufficient statistic for the worker's employment probability and evolves according to the differential equation.

In general, solving for an equilibrium involves solving for a fixed point of functional equation. However, given the assumptions on our market clearing process, the equilibrium exhibits a very simple form. First of all, we focus on the worker's optimization problem. Since the worker is choosing from a discrete number of applications, looking at the marginal benefit of an additional application suffices. The marginal benefit of the $k+1$-st application, denoted MB_{k+1}, given $\frac{A}{M} \in (0,1)$ is:

$$MB_{k+1}(\tfrac{A}{M}) = 1 - \Big(\frac{A}{M}\Big)^{k+1} - \Big(1 - \Big(\frac{A}{M}\Big)^{k}\Big) = \Big(\frac{A}{M}\Big)^{k}\Big(1 - \frac{A}{M}\Big).$$

The marginal benefit of first five applications as functions of $A(x)$ are given in Figure 1. We first note that, for a fixed $A(x)$, the marginal benefit always decreases for higher numbers of applications. With constant application c, a worker with success probability $\frac{A(x)}{M}$ chooses the largest integer $k(x)$ such that $MB_{k(x)} \geq c$.

For later references, it is useful to denote as $K(\frac{A}{M})$ the solution to the optimal application problem when the probability of success is $\frac{A}{M}$. Then the second equilibrium condition can be written as

$$k(x) \in K(\frac{A(x)}{M}). \tag{2}$$

Secondly, we note that the marginal benefit of the first application ($MB1$) is a monotone decreasing function of A, while the other marginal benefits are single peaked in A. Given this observation, and from the continuity and monotonicity of function $A : [0, 1] \to \mathbb{R}_+$, following from continuity of types, we see that:

Theorem 1. The equilibrium number of applications $k(x)$ is single-peaked in x.

Proof. (sketch) The marginal benefit is single-peaked in $A(x)$ and decreasing in k. Agents increase number of applications as long as marginal benefit exceeds marginal cost. If the cost curve cuts through the marginal benefit curve of the kth application it must have cut through $(k-1)$th for smaller $A(x)$ and through $(k-2)$th for even smaller $A(x)$ and so on. Similarly it must cut through either the marginal benefit curve for the $(k+1)$th application for larger $A(x)$ or again cut through the kth and so on (see Figure 1). Since $A(x)$ is strictly decreasing in x the result immediately follows. □

These observations allow us to restrict the set of functions that are feasible as a solution. Also, $A(x)$ is continuous, since the accumulation increases gradually by type, but not differentiable at types where the number of applications $k(x)$ changes. Hence, it is "pieced together" by differentiable parts that are connected at cutoff types. For function k, there can be multiplicity. First, note that any types $x \leq \underline{x}$, where $A(\underline{x}) = M(1-c)$, are indifferent between applying and not applying since with $\frac{A}{M} = 1 - c$, both 0 and 1 are best responses. Without further restriction, any k with measure zero subset of these agents ($x \leq \underline{x}$) sending out one application would be an equilibrium. Furthermore a finite number of cutoff types are indifferent between two adjacent applications. Requiring left- or right-continuity of the decision rule helps us to eliminate trivial multiple equilibria.

Normalization Note that since firms base their hiring decision only on rank and not on the cardinality of types, the equilibrium outcome is invariant to the worker type distribution. The following proposition shows that we can without loss of generality restrict attention of our analysis to a uniform F.

Proposition 1. We can equivalently express the equilibrium with underlying distribution F by functions $(\tilde{A}, \tilde{K})$ of the transformed type $\tilde{x} = F(x)$ satisfying the following equations:

$$-\tilde{A}'(\tilde{x}) = 1 - \left(\frac{\tilde{A}(\tilde{x})}{M}\right)^{\tilde{K}(\tilde{x})}$$

$$\tilde{K}(\tilde{x}) = \arg\max_k 1 - \left(\frac{\tilde{A}(\tilde{x})}{M}\right)^k - ck$$

Proof. We change variables and define $\tilde{x} = F(x)$, and $F^{-1}(\tilde{x}) = x$. Assuming F is continuous, differentiable and one-to-one, its inverse is well-defined and the equilibrium conditions can be rewritten as

$$-A'(F^{-1}(\tilde{x})) = F'(F^{-1}(\tilde{x}))(1 - \left(\frac{A(F^{-1}(\tilde{x}))}{M}\right)^{k(F^{-1}(\tilde{x}))}),$$

$$k(F^{-1}(\tilde{x})) = K(A(F^{-1}(\tilde{x})))$$

$F'(F^{-1}(\tilde{x})) = \frac{1}{\frac{dF^{-1}(\tilde{x})}{d\tilde{x}}}$, so that

$$-A'(F^{-1}(\tilde{x}))\frac{dF^{-1}(\tilde{x})}{d\tilde{x}} = 1 - \left(\frac{A(F^{-1}(\tilde{x}))}{M}\right)^{k(F^{-1}(\tilde{x}))}.$$

Let $A(F^{-1}(\tilde{x})) = \tilde{A}(\tilde{x})$, and $k(F^{-1}(\tilde{x})) = \tilde{K}(\tilde{x})$. Then:

$$-\tilde{A}'(\tilde{x}) = 1 - \left(\frac{\tilde{A}(\tilde{x})}{M}\right)^{\tilde{K}(\tilde{x})},$$

$$\tilde{K}(\tilde{x}) \in K(\frac{\tilde{A}(\tilde{x})}{M})$$

□

Corollary 1. For any two distributions F, G, and a pair $(x, x') \in [0,1]^2$ such that $F(x) = G(x')$, in equilibria, $A^F(x) = A^G(x')$.

Clearly, given $\tilde{A}(F(x)) = \tilde{A}(G(x'))$, regardless of the distribution, workers of the same rank face the same employment probability. This is intuitive since the probability is only determined by the failure to coordinate with higher types. Any given mass of higher type workers results in the same mass of employed higher type workers and leads to a corresponding reduction in employment probability. The following result is also immediate:

Corollary 2. Total employment does not depend on F.

Proof. Total employment in the model is $A(F(0)) = \tilde{A}(0)$. □

For notational simplicity, hereafter, we rewrite the type $\tilde{x}$ as x, and the equilibrium objects $(\tilde{A}, \tilde{K})$ as (A, K).

Corollary 3. If type zero workers send out zero applications then total employment $A(0)$ is given by $M(1-c)$. If type zero workers send out a strictly positive number of applications, total employment is strictly less than $M(1-c)$.

Proof. The above establishes that if $A(x') = M(1-c)$ for some x', then all lower types must send out zero applications. Hence, employment can never exceed $M(1-c)$. If some workers send out zero applications it must be that employment has reached $M(1-c)$. If all types apply strictly positive number of times, it must be that the threshold was not yet reached, i.e., $A(0) < M(1-c)$. □

4 Firm Entry

In this section we examine the connection between application behaviour and the size of the vacancy pool M, e.g. in the case of business cycle fluctuations and firm entry. It is not clear whether workers exert more or less search effort for different M. Fewer jobs might urge workers to try harder to secure a job or make workers more inclined to give up.

Figure 2 depicts the difference in marginal benefits for two sizes of job pools, $M = 1$ and $M' = 0.8$, represented by the blue and red graph respectively. The marginal benefit of the first application $1 - \frac{A(x)}{M}$ decreases for all levels of A, but the marginal benefit of additional applications increase for small A, and decrease for large A, i.e. the blue line lies above the red for some range of $A(x)$ and vice versa. Since $A(x)$ is a monotone function of x, this suggests that for fewer available jobs higher type workers might increase their number of applications while lower types might decrease their number of applications. We investigate this and provide proofs in the following section.

We use subscripts 0 and 1 to denote variables given M_0 and M_1 with $M_0, M_1 > 0$. Define $\gamma = \frac{M_1}{M_0}$.

Proposition 2. Denote tentative[1] cutoff types that satisfy

$$\left(\frac{A(x)}{M_i}\right)^j \left(1 - \frac{A(x)}{M_i}\right) = c, \quad i \in \{0,1\},\ j \in \mathbb{Z}_+$$

by x_i^{jh} for the higher type and x_i^{jl} for the lower type. (There must exist two since the

[1]These types need not be greater or equal to zero and hence might only be hypothetical types that no actual worker has.

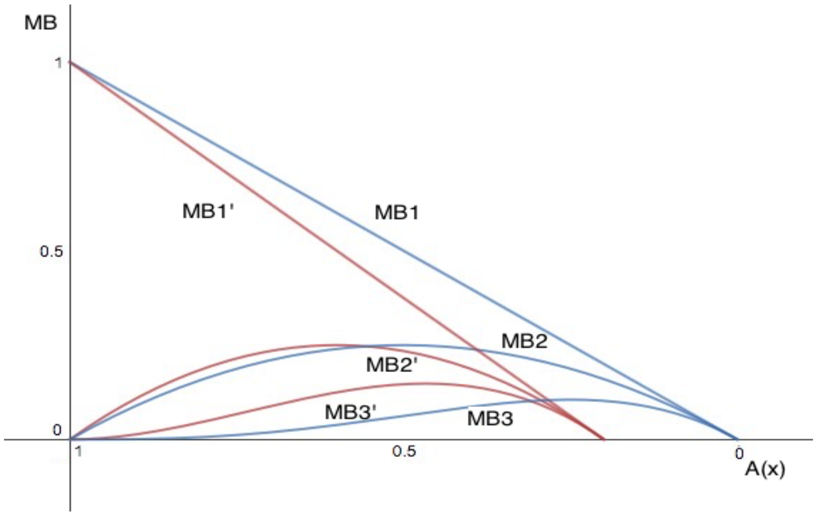

Figure 2: Marginal Benefit for Applications 1-3

marginal benefit curves are hump-shaped). We suppress the second superscript for convenience when not needed. These cutoff types are indifferent between making j applications and $j+1$ applications. Then for all x_i^j, $\gamma = \frac{A(x_1^j)}{A(x_0^j)}$ and $\gamma(1-x_0^j) = 1 - x_1^j$. That is, the cutoff accumulation and mass of workers with types larger than x^j are proportional to M.

Proof. It follows directly from the equation that the cutoff accumulation is proportional to M.

For the second part, define $\hat{A}(\frac{1-x}{M}) = \frac{A(x)}{M}$. Note that

$$\frac{d}{dx}\hat{A}(\tfrac{1-x}{M}) = -\hat{A}'(\tfrac{1-x}{M})\frac{1}{M} = -\frac{A'(x)}{M}.$$

The equilibrium can be restated with the new transformed variable $\hat{x} = \frac{1-x}{M}$ and equilibrium functions $(\hat{A}, \hat{K})$, $\hat{A} : [0, \frac{1}{M}] \to \mathbb{R}_+$, and $\hat{K} : [0, \frac{1}{M}] \to \mathbb{Z}_+$:

$$\hat{A}'(\hat{x}) = 1 - \hat{A}(\hat{x})^{\hat{K}(\hat{x})}, \quad \hat{A}(0) = 0$$
$$\hat{K}(\hat{x}) \in K(\hat{A}(\hat{x}))$$

These equilibrium objects do not depend on M except through changes of the domain. At the cutoffs we have

$$\hat{A}(\hat{x})^j \left(1 - \hat{A}(\hat{x})\right) = c, \ j \in \mathbb{Z}_+$$

Since $\hat{A}$ is strictly increasing in $\hat{x}$ up until the last cutoff and as we can see from the above equation the cutoff values of $\hat{A}$ are invariant in M, the transformed tentative cutoff types $\hat{x_i}^j = \frac{1-x_i^j}{M}$ are also invariant in M. Hence we have

$$\hat{x}_0^j = \frac{1-x_0^j}{M_0} = \frac{1-x_1^j}{M_1} = \hat{x}_1^j, \quad \forall j,$$

and therefore $\gamma(1-x_0^j) = 1-x_1^j$. □

We also know that $A(x)$ changes proportionally in M for all x. Since $\hat{A}(\hat{x})$ is invariant to changes in M, for any level of accumulation A, $\gamma(1-A_0^{-1}(A)) = (1-A_1^{-1}(A))$. Graphically $K(F(x))$ and $A(F(x))$ "contract" proportionally towards the top of the type distribution. We see that a change in M does not change the number of tentative cutoff types or the employment probability of cutoff types.

The following two propositions describe how total employment depends on M.

Proposition 3. If $K_0(0) = 0$ and $K_1(0) = 0$, then the ratio of total employment is proportional to M_0 over M_1. to M, that is, $A_1(0) = \gamma A_0(0)$. Hence, the hiring probability for firms remains the same.

Proof. Follows from Corollary 3. □

For cases where none of the workers sent out zero applications we derive the following result:

Proposition 4. If $K_0(0) > 0$, The difference in total employment is less than proportional to the difference in vacancy pool size, that is, given $\gamma > 1$, then $\gamma A_0(0) > A_1(0)$, and if $\gamma < 1$, then $\gamma A_0(0) < A_1(0)$. Therefore the hiring probability decreases for an increase in M.

Proof. See Appendix A.4.1. □

Denote the maximum number of applications $\bar{k}$. Then the lower of the two cutoffs type switching between $\bar{k}$ and $\bar{k}-1$ applications is $x_i^{\bar{k}-1l}$ and the higher $x_i^{\bar{k}-1h}$ for $i = 0, 1$. We establish the following for application behaviour across types:

Proposition 5. All worker types higher than $x_1^{\bar{k}-1l}$ apply weakly more and all lower worker types apply weakly less with M_0 compared to M_1 for $\gamma > 1$.

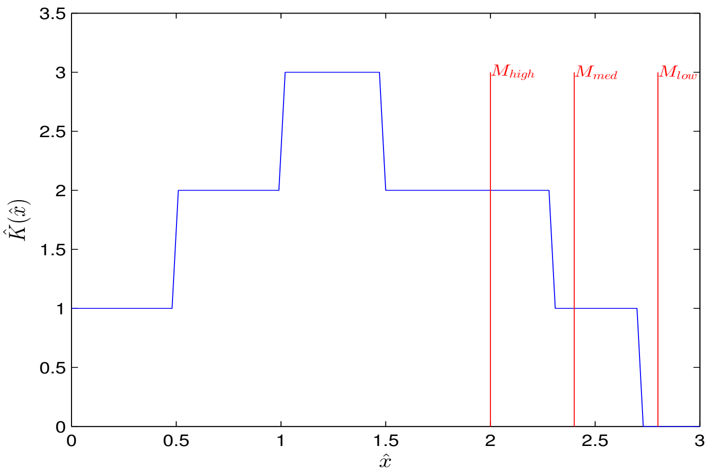

Figure 3: Observed investment behavior with varying M

Proof. We know that for all cut-off types x_i^j,

$$\frac{1 - x_0^j}{M_0} = \frac{1 - x_1^j}{M_1}$$

hence, $x_1^j < x_0^j$ if $\gamma < 1$. This, together with the fact that with M_1 jobs the number of applications weakly increases in type for all types lower than $x_1^{\bar{k}-1l}$, results in these types sending out fewer or just as many applications as with M_0. On the other hand, the number of applications weakly decreases in type for all types higher than $x_1^{\bar{k}}l$. Since the cutoff types are lower in M_1 than in M_0 they apply to at least as many positions. A similar argument shows that for $\gamma > 1$ the threshold type is $\bar{x}_1^{\bar{k}}$. □

Note that worker types who apply more with a smaller vacancy pool size compared to a larger pool size do not need to have a higher employment probability, since the fraction of available jobs not taken by higher types is also smaller.

However, given that the difference in the pool sizes is sufficiently small, we can find an $\epsilon > 0$, such that workers with types greater than the cutoff type minus ϵ (close to the cutoff) who apply more with a smaller pool have a higher employment probability. To see why, note that the employment probability jumps discontinuously when the number of applications is increased.

Proposition 6. The maximum number of applications observed $\bar{K}$ is weakly increasing if $\gamma < 1$, but is weakly decreasing if $\gamma > 1$.

Proof. Figure 3, which plots an example of $(\hat{x}, K(\hat{x}))$ illustrates what happens when M

changes. $\hat{x} = 0$ is the highest type $x = F(1)$, and the type decreases as we move right on the x-axis. As shown before $\hat{K}$ and $\hat{x}$ are invariant to M. We plot the lowest type, $\hat{x} = \frac{1}{M}$, and his number of application for different M. A high M corresponds to a smaller domain and the observed application pattern being a smaller cutout from the whole picture; while low M corresponds a large domain and also a larger cutout. In the limit, as M becomes very large, everyone applies only once; as M becomes very small, most people apply 0, while a small group of people at the top keeps the single-peaked application pattern. □

4.1 Worker quality and Free Entry

Given that we understand now how application behaviour changes for a different size of the vacancy pool, we analyze now the difference in average quality of a hired worker. We find an increasing relationship which means that in the case of free entry of firms and a cost $v > 0$ of posting a vacancy, there is a unique M such that there is no more exit and entry.

First, we define the measure of 'quality' of workers hired.

Definition 2. Expected quality of a worker hired by the firm:

$$E[x|x \text{ hired}] = \frac{\int_{\underline{x}}^{1} -A'(x)x\,dx}{M} = \int_{\underline{x}}^{1} -\frac{A'(x)}{M} x\,dx.$$

An economy (M, F) is identified with the measure of firms, M, and the distribution of worker types, F. Using integration by parts, we obtain the following sufficient condition for comparing expected qualities.

Proposition 7. For two economies, (M_1, F_1) and (M_2, F_2), if $\frac{A_1(x)}{M_1} \leq \frac{A_2(x)}{M_2}$ for all x, then the expected quality of hired workers in economy 2 is weakly greater than of those in economy 1.

Proof. Since the endpoints are the same: $A_1(1) = A_2(1) = 0$,

$$\frac{A_1(x)}{M_1} \leq \frac{A_2(x)}{M_2} \iff 0 \leq \int_{x}^{1} -\frac{A_1'(y)}{M_1} dy \leq \int_{x}^{1} -\frac{A_2'(y)}{M_2} dy$$

Integration by parts: $\forall x$,

$$\begin{aligned}\int_x^1 -\frac{A_1'(y)}{M_1} y\, dy &= \frac{A_1(x)}{M_1}x + \int_x^1 \frac{A_1(y)}{M_1}\, dy \\ &\leq \frac{A_2(x)}{M_2}x + \int_x^1 \frac{A_2(y)}{M_2}\, dy \\ &= \int_x^1 -\frac{A_2'(y)}{M_2} y\, dy\end{aligned}$$

□

Note that this implies that for two economies that are identical except that the worker type distribution in one first-order stochastically dominates the other, $F_1(x) \leq F_2(x)$ for all x, then the expected hired worker quality in the second economy will be higher in the first, since $\frac{A_1(x)}{M} \leq \frac{A_2(x)}{M}$. Define the *market tightness* up until type x worker, of an economy (M, F), as

$$T(x) = \frac{1 - F(x)}{M}.$$

This is the ratio of measure of workers above type x and the measure of firms. The next proposition shows that the stochastic dominance relation in Proposition 7 is implied by the dominance relation between the T's.

Proposition 8. For arbitrary x, $T_1(x) = \frac{1-F_1(x)}{M_1} \leq \frac{1-F_2(x)}{M_2} = T_2(x)$ implies that $\frac{A_1(x)}{M_1} \leq \frac{A_2(x)}{M_2}$.

Proof.

$$\begin{aligned}\frac{A_1(x)}{M_1} \leq \frac{A_2(x)}{M_2} \quad &\iff \quad \int_x^1 -\frac{A_1'(y)}{M_1}\, dy \leq \int_x^1 -\frac{A_2'(y)}{M_2}\, dy \\ \iff \quad \int_x^1 (1 - \Big(\frac{A_1(y)}{M_1}\Big)^{k_1(y)}) \frac{f_1(y)}{M_1}\, dy &\leq \int_x^1 (1 - \Big(\frac{A_2(y)}{M_2}\Big)^{k_2(y)}) \frac{f_2(y)}{M_2}\, dy \\ \iff \quad \int_0^{\frac{1-F_1(x)}{M_1}} (1 - \tilde{A}(z)^{\tilde{K}(z)}) dz &\leq \int_0^{\frac{1-F_2(x)}{M_2}} (1 - \tilde{A}(z)^{\tilde{K}(z)}) dz\end{aligned}$$

Last line by the change of variables $z = \frac{1-F_i(x)}{M_i}$. □

As an immediate corollary, we observe that:

Corollary 4. More firms in the economy imply a lower expected quality of workers hired.

Proof. $M_1 > M_2$, then $\frac{A_1(x)}{M_1} \leq \frac{A_2(x)}{M_2}$ for all x. □

It immediately follows that given a worker type distribution and entry cost $v > 0$ there is an unique M that solves the free entry condition:

$$E[x|x \text{ hired}] - v = 0$$

5 Application Costs

Now we would like to examine the relationship between application behaviour and cost c. Note that a higher c is in fact equivalent to a lower wage since the application decision is determined by the condition $MB \geq c$ or:

$$w\Big(\frac{A(x)}{M}\Big)^k(1 - \frac{A(x)}{M}) \geq c$$

We see that the decision is only dependent on the ratio of cost to wage. We now proceed to contrast utility for each worker type in case of a high and a low cost.

When the cost of application c is low, an individual worker's utility is affected in two ways: on one hand net the return of each application is high for this worker and on the other hand this is also true for higher ranked workers, who apply more. The first effect is positive while the second effect is negative. This leads the counter-intuitive result that some workers might in fact be worse off with lower c. The utility for an individual who faces $(A, c) \in (0, 1 - c) \times (0, 1)$ is given by

$$v(A, c) = \max_{k \in \mathbb{Z}_+} 1 - A^k - ck$$

Denote the optimal application decision as $k^*(A, c)$. Generally, $A^{k^*(A,c)}(1 - A) < c$ and $A^{k^*(A,c)-1}(1 - A) > c$ and the derivative exists. Hence, by the envelope theorem:

$$\frac{\partial v}{\partial c} = -k^*(A, c).$$

The derivative with respect to A if the optimal application number does not change is

$$\frac{\partial v}{\partial A} = -k^*(A, c)A^{k^*(A,c)-1}$$

Hence, the overall effect when $k^*(A, c)$ remains constant is

$$\begin{aligned}\frac{dv}{dc} &= \frac{\partial v}{\partial A}\frac{dA}{dc} + \frac{\partial v}{\partial c} \\ &= -kA^{k-1}\frac{dA}{dc} - k\end{aligned}$$

Note that $\frac{dA}{dc} < 0$. Therefore, if $\left|A^{k-1}\frac{dA}{dc}\right| > 1$, the agent is better off with higher c. This is more likely to be the case for low worker types who face a high A and also $\frac{dA}{dc}$. This can be seen in Figure 4 where we plot the worker utility for different c and $M = 0.6$. Utility for types around 0.4 is higher for $c = 0.248$ than $c = 0.148$.

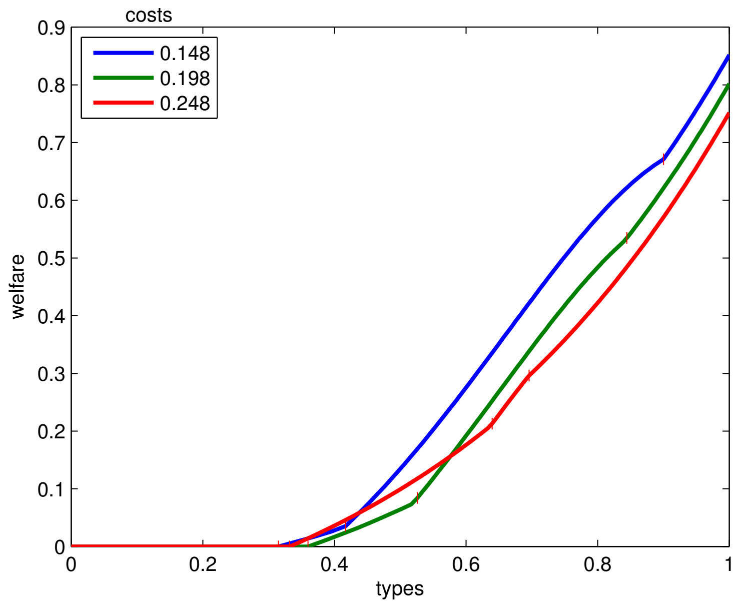

Figure 4: Welfare change as c varies

Another way to illustrate this is to consider $c \approx 0$. Note that in a frictionless matching market, the unique stable matching outcome will be perfectly assortative: All workers in the top M-percentile find a job. If $M < 1$ then workers of types $x < 1 - M$ will have zero probability of being hired. Now turn to a case with a slightly larger cost. Since the equilibrium varies continuously with the parameters, we conjecture that up to top M-th percentile of workers get hired with probability slightly less than one and workers with types $x < 1 - M$ now have a strictly positive employment probability. Hence, their situation has improved and this shows that some types might benefit from higher cost.

6 Efficiency

We now discuss efficiency in our model. Note that each worker's search imposes a negative externality on lower type workers, since they 'take away' jobs from lower types. When evaluating whether to make an additional application, workers weigh the increase in employment probability against c but do not consider whether a newly found job would have otherwise been taken by another lower type worker.

To illustrate the discrepancy between individual and social welfare, we can consider the problem of a social planner whose objective it is to maximize total worker utility, also equal to total employment minus total application costs.

Theorem 2. The planner's solution has the number of applications be monotonically decreasing in worker type. Furthermore there are worker types sending out more than one application only if all workers apply at least once.

Proof. We provide detailed proofs in A.3. The idea is as follows, the planner does not care about which workers are employed but only about total employment. If some lower type workers sent out fewer applications than some higher type workers it would always result in higher total employment to switch the numbers of applications for the two groups, while keeping application costs constant. □

This result must not apply if the planner does care about worker types. However, our formulation of the planner's objective can be seen as the limit case when worker types are arbitrarily close to each other.

7 Conclusion

We examined how worker rank heterogeneity and inability to coordinate translate into different search effort and matching outcomes for different types. If a worker's rank is sufficiently high he is encouraged to apply more than slightly higher rank workers to insure against the event of firms choosing someone else. However if a worker's rank is sufficiently low, his poor chances of success per-application might discourage him from applying as much as his higher rank competitors. In this case the lower per-application success rate and the lower search effort put the worker at an even greater disadvantage. This offers a novel insight to unequal labour outcomes. Higher unemployment among low-skilled workers must not be driven by the demand side but can also arise due to endogenous adjustments on labour supply.

Furthermore, we examined how our results depend to the size of the vacancy pool and application costs. We conclude by discussing that worker welfare might be improved by increasing application costs due to negative search externalities.

References

Albrecht, James, Pieter A Gautier, and Susan Vroman. 2006. "Equilibrium directed search with multiple applications." *The Review of Economic Studies*, 73(4): 869–891.

Albrecht, James, Serene Tan, Pieter Gautier, and Susan Vroman. 2004. "Matching with multiple applications revisited." *Economics Letters*, 84(3): 311–314.

Albrecht, James W, Pieter A Gautier, and Susan B Vroman. 2003. "Matching with multiple applications." *Economics Letters*, 78(1): 67–70.

Azevedo, Eduardo M, and Jacob D Leshno. 2014. "A supply and demand framework for two-sided matching markets." *Available at SSRN 2260567.*

Burdett, Kenneth, and Melvyn G Coles. 1999. "Long-term partnership formation: Marriage and employment." *The Economic Journal*, 109(456): 307–334.

Chade, Hector. 2001. "Two-sided search and perfect segregation with fixed search costs." *Mathematical Social Sciences*, 42(1): 31–51.

Chade, Hector, and Lones Smith. 2006. "Simultaneous search." *Econometrica*, 74(5): 1293–1307.

Chade, Hector, Gregory Lewis, and Lones Smith. 2014. "Student portfolios and the college admissions problem." *The Review of Economic Studies*, 81(3): 971–1002.

Galenianos, Manolis, and Philipp Kircher. 2009. "Directed search with multiple job applications." *Journal of Economic Theory*, 144(2): 445–471.

Hornstein, Andreas, and Marianna Kudlyak. 2015. "Estimating Matching Effi ciency with Variable Search Effort." mimeo.

Kircher, Philipp. 2009. "Efficiency of simultaneous search." *Journal of Political Economy*, 117(5): 861–913.

Petrongolo, Barbara, and Christopher A Pissarides. 2001. "Looking into the black box: A survey of the matching function." *Journal of Economic literature*, 39(2): 390–431.

Shimer, Robert. 2004. "Search intensity." *Unpublished paper.*

Shimer, Robert, and Lones Smith. 2000. "Assortative matching and search." *Econometrica*, 68(2): 343–369.

add stigler 1961 for discrete Pissarides C A 2000 Equilibrium Unemployment Theory. 2nd ed. MIT Press, Cam- bridge, MA for continuous model

A Appendix

A.1 Derivation of the DE

Consider an economy where the worker types are distributed according to density function $f : [0,1] \to \mathbb{R}$, but the firms perceive only a coarse measure of types: firms sort workers in N equal-sized groups according to type. Suppose also that all workers in the same group can coordinate their applications, but applications from different groups of workers are not coordinated. The top group of workers will always get the job and will take up $\frac{1}{N}$ out of M jobs. An application from second tier workers lands on a vacant job with probability $\frac{M-\frac{1}{N}}{M}$ and overlaps with an application from first tier workers with probability $\frac{1}{NM}$. In total $\frac{1}{N}(1-\frac{1}{NM})$ workers from the second tier find a job. In general, let A_k be the mass of workers up up to tier k who are matched; $A_{k+1} - A_k$ is the mass of $k+1$-th tier workers who found jobs. Then,

$$A_{k+1} - A_k = \frac{1}{N}(1 - \frac{A_k}{M}).$$

A worker with type $x < 1$ is located at $\lceil N(1-F(x)) \rceil$-th tier. Reindexing the types:

$$\frac{A_{(1-F(x))+\frac{1}{N}} - A_{(1-F(x))}}{1/N} = 1 - \frac{A_{(1-F(x))}}{M}$$

By letting $N \to \infty$, the difference equation converges to a differential equation:

$$A'(1-F(x)) = 1 - \frac{A(1-F(x))}{M}$$

Defining $\tilde{A}(x) = A(1-F(x))$, the equation is equivalently written as

$$-\frac{\tilde{A}'(x)}{f(x)} = 1 - \frac{\tilde{A}(x)}{M}$$

A.2 Proof of the DE

Let $\varphi(x) \in [0,1]$ be the probability that type x's application results in an offer from firm. An offer from the firm means that all types strictly above x has chosen which firm to go, and $\varphi(x)$ is the probability that an x type's application is not taken up by types above x.

Assume that application behavior $K : [0,1] \to \mathbb{Z}_+$ which maps types to applications, is piecewise left-continuous. Then for all types x who apply $k = K(x)$ times, there exists a finite number of cutoffs $x < x_1 < x_2 < \ldots < x_{N-1} < x_N = 1$ such that all types in (x, x_1) apply k times and any types in (x_i, x_{i+1}) send out same number of applications. We refer to the types in (x_i, x_{i+1}) as the i-th group. Denote by $a_i \in \mathbb{Z}_+$ the applications sent out by group i, and $\lambda_i = F(x_{i+1}) - F(x_i)$ the mass of group i. For the collection $(a_i, \lambda_i)_{i=1}^N$, the Poisson arrival rate of applications from types higher than x is given by $\frac{\lambda(x)}{M}$, where

$$\lambda(x) = \sum_i a_i \lambda_i + k(F(x_1) - F(x)))$$

An application by x is successful when either (1) there are no competing applications or (2) there are competing applications but all competitors choose other offers. Conditioning on arrival, probability that it is from i-th group is $\frac{a_i \lambda_i}{\lambda}$. The competitor $z_i \in (x_i, x_{i+1})$ randomizes equally over all the offers if there are multiple. The probability that z_i picks offers other than the one which overlaps with x is given by:

$$P_i(z_i;\varphi) = \sum_{j=0}^{a_i-1} \frac{j}{j+1} \binom{a_i-1}{j} \varphi(z_i)^j (1-\varphi(z_i))^{a_i-1-j}.$$

Since this is true for any $z_i \in (x_i, x_{i+1})$, given that the one application came from this group, the conditional probability that the group does not take the job is given by

$$\int_{x_i}^{x_{i+1}} P_i(z;\varphi) \frac{f(z)}{\lambda_i} dz := p_i(\varphi)$$

Therefore, overall, in expectations, the probability that the job is not taken by any group is given by

$$J(x) = \sum_i \frac{a_i \lambda_i}{\lambda(x)} p_i(\varphi) + \frac{k}{\lambda(x)} \int_x^{\hat{x}} P_k(z;\varphi) f(z) dz$$

The probability is independent for multiple number of overlaps, and with $n > 1$ overlaps, the probability that the job is not taken is simply $J(x)^n$.

Hence, by the definition of φ:

$$\varphi(x) = \sum_{k=0}^{\infty} \frac{(\frac{\lambda(x)}{M})^k e^{-\frac{\lambda(x)}{M}}}{k!} J(x)^k = e^{\lambda(x)(J(x)-1)}$$

The last equality follows from the fact that $E[a^X] = \exp(\lambda(a-1))$ for $X \sim Pois(\lambda)$ and $a > 0$. Since everyone in the neighborhood of x also applies k times, taking log and differentiating with respect to x yields:

$$\frac{\varphi'(x)}{\varphi(x)} = \frac{d}{dx}(\frac{\lambda(x)}{M}(J(x)-1))$$

Note that $\lambda(x)J(x) = \sum_i a_i \lambda_i p_i(\varphi) + k\int_x^{\hat{x}} P_k(z;\varphi) f(z)dz$, so that

$$\frac{d}{dx}(\frac{\lambda(x)}{M}J(x)) = -kP_k(x;\varphi)\frac{f(x)}{M},$$

and $\frac{d}{dx}\big(\lambda(x)\big) = -kf(x)$. Hence,

$$\frac{\varphi'(x)}{\varphi(x)} = -k\frac{f(x)}{M}(P_k(x;\varphi)-1)$$

$$\begin{aligned} P_k(x;\varphi) &= \sum_{j=0}^{k-1} \frac{j}{j+1}\binom{k-1}{k-1-j}\varphi(x)^j(1-\varphi(x))^{k-1-j} \\ &= 1 - \sum_{j=0}^{k-1} \frac{1}{j+1}\binom{k-1}{k-1-j}\varphi(x)^j(1-\varphi(x))^{k-1-j} \end{aligned}$$

Therefore,

$$\begin{aligned} \varphi(x)(kP_k(x;\varphi)-k) &= -\sum_{j=0}^{k-1} \frac{k}{j+1}\binom{k-1}{k-1-j}\varphi(x)^{j+1}(1-\varphi(x))^{k-1-j} \\ &= -\sum_{l=1}^{k}\binom{k}{l}\varphi(x)^l(1-\varphi(x))^{k-l} \\ &= -(1-(1-\varphi(x))^k) \end{aligned}$$

In conclusion:

$$\varphi'(x) = \frac{f(x)}{M}(1-(1-\varphi(x))^k))$$

With no coordination among workers, Appendix A.1 shows that $1-\varphi(x)$ the proportion of jobs that are already taken up by types higher than x. That is, $1-\varphi(x)=\frac{A(x)}{M}$, and

$$-A'(x) = f(x)(1-(\frac{A(x)}{M})^k).$$

A.3 Numerical Example with Two Applications

Threshold types In this section we present the analytical solutions for $\frac{4}{27} \leq c < \frac{1}{4}$, the case in which at most two applications are sent out and there is a positive measure of worker types sending out two applications.

For the cutoff type switching from one to two applications $\hat{x}^{1h}$ it holds that:

$$MB(\hat{x}^{1h}) = c$$

$$(1-e^{-\hat{x}^{1h}})e^{-\hat{x}^{1h}} = c$$

Hence:

$$\hat{x}^{1h} = -ln\left(\frac{1}{2} + \sqrt{\frac{1}{4} - c}\right)$$

Define $\beta = \frac{1}{2} + \sqrt{\frac{1}{4} - c}$ then $\hat{x}^{1h} = -ln(\beta)$. The form of the accumulation function when agents make two applications is:

$$A(x) = 1 - \frac{2}{e^{2x+s}+1}$$

And at $\hat{x}^{1h}$ it must be that

$$1 - e^{-\hat{x}^{1h}} = 1 - \frac{2}{e^{2\hat{x}^{1h}+s}+1}$$

and

$$e^s = 2e^{-\hat{x}^{1h}} - e^{-2\hat{x}^{1h}}$$

$$e^s = 2\beta - \beta^2$$

Then for the cutoff type switching back from two to one application $\hat{x}^{1l}$ we have:

$$MB(\hat{x}^{1l}) = c$$

$$\left(1 - \frac{2}{e^{2\hat{x}^{1l}+s}+1}\right)\frac{2}{e^{2\hat{x}^{1l}+s}+1} = c$$

Define $\delta = \frac{1}{2} - \sqrt{\frac{1}{4} - c}$ then:

$$\frac{2}{e^{2\hat{x}^{1l}+s}+1} = \frac{1}{2} - \sqrt{\frac{1}{4} - c} = \delta$$

$$\hat{x}^{1l} = \frac{1}{2}ln\left(\frac{2-\delta}{\delta}e^{s}\right) = \frac{1}{2}ln\left(\frac{2-\delta}{\delta\beta(2-\beta)}\right)$$

For the accumulation it must hold that:

$$1 - \frac{2}{e^{2\hat{x}^{1l}+s}+1} = 1 - e^{-\hat{x}^{1l}+t}$$

$$\delta = e^{-\hat{x}^{1l}+t}$$

$$e^{t} = \delta e^{\hat{x}^{1l}} = \delta\sqrt{\frac{2-\delta}{\delta\beta(2-\beta)}} = \sqrt{\frac{\delta(2-\delta)}{\beta(2-\beta)}}$$

Find the cutoff type switching from one to zero applications $\hat{x}^{0}$:

$$e^{-\hat{x}^{0}+t} = c$$

$$\hat{x}^{0} = -ln(c) + t = -ln(c) + ln(\delta) + \hat{x}^{1l}$$

Note that $\delta\beta = c$, hence:

$$\hat{x}^{0} = ln(\frac{1}{\beta}) + \hat{x}^{1l} = \hat{x}^{1h} + \hat{x}^{1l}$$

Hence, last cutoff is the sum of the two previous cutoffs. It is clear that $\frac{\partial\hat{x}^{1h}}{\partial c} > 0$. The derivative of the other cutoff with respect to cost is:

$$\frac{\partial\hat{x}^{1l}}{\partial c} = \frac{1}{2}\frac{\delta\beta(2-\beta)}{2-\delta}\frac{(c)(2-\beta)(-\delta') - (2-\delta)((2-\beta)-\beta'(c))}{c^{2}(2-\beta)^{2}}$$

$$= \frac{1}{2}\frac{(c)(2-\beta)(-\delta') - (2+c) + (2-\delta)\beta'(c))}{c(2-\beta)(2-\delta)}$$

$$= \frac{1}{2}\frac{c\left(-\frac{3}{2\sqrt{\frac{1}{4}-c}}\right) - (2+c)}{c(2+c)} < 0$$

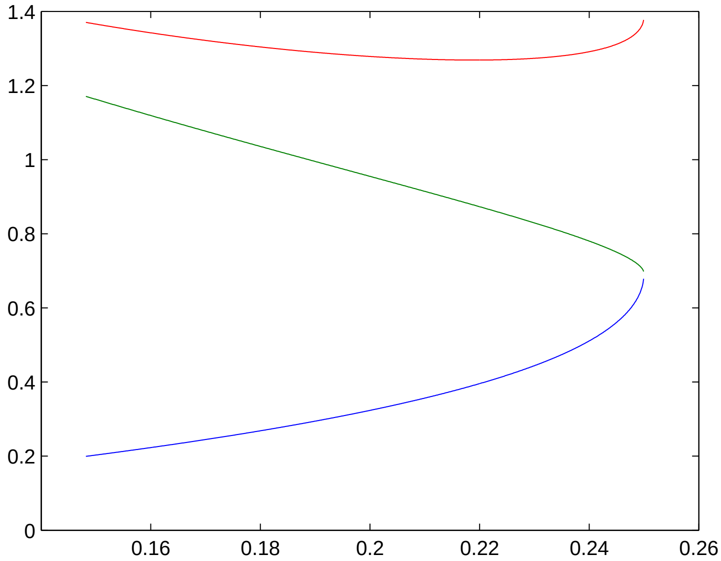

Figure 5: c and threshold types

So the range of worker types applying twice shrinks with increasing cost. Since the third cutoff is the sum of the two cutoffs:

$$\frac{\partial \hat{x}^0}{\partial c} = \frac{\partial \hat{x}^{1l}}{\partial c} + \frac{\partial \hat{x}^{1h}}{\partial c}$$

The direction of changes in $\hat{x}_0$ is indeterminate. Figure A.3 plots change of cutoff types $(\hat{x}_0, \hat{x}^{1l}, \hat{x}^{1h})$ as c varies from $\frac{4}{27}$ to $\frac{1}{4}$. Note that $\hat{x}_0$ (top graph) does not vary monotonically with c.

Social Planner Problem Our welfare objective is to maximize total worker utility, equal to total employment minus total application cost. This is appropriate if the main interest is to reduce unemployment, perhaps in conjunction with equity and crime concerns, and the skills of employed workers do not have a large enough effect on output in the economy.

Given an exogeneous cost c the social planner's problem is to maximize welfare as a function of equilibrium objects (A, K) through choosing post-tax cost c' faced by the applicants. Denote the equilibrium with cost c' to be $(A_{c'}, K_{c'})$ then the problem is:

$$\max_{c'} W_{c'} = A_{c'}(0) - c \int_0^1 K_{c'}(z) dz.$$

We define $t = c' - c$ as the tax/subsidy per application.

In the following analysis we again restrict ourselves to cases in which $c \geq \frac{4}{27}$, that is, workers send out at most two applications, and provide some intuition why similar results should hold true for larger numbers of application.

Let us first consider $c > \frac{1}{4}$, such that all workers send out at most one application. If all workers apply exactly once, imposing a tax will only have an effect on welfare if it is large enough so that some low type workers do not apply. If there already exist workers at the tail of the type distribution who do not apply, a tax will shift this cutoff upward discouraging even more workers from applying.

We can see that for this range of cost a tax can never be welfare-improving. Without taxation every worker who chooses to apply has an employment probability at least as large as c. By imposing a tax, a range of worker types lowest in the applicant pool will be discouraged from applying. These workers do not impose any externalities on others except on each other. Changing their application behaviour does not benefit anyone since all lower types do not apply and increasing the cost will not change that. Hence, any tax, if it has an effect, decreases total application cost to a smaller extent than it decreases total employment and hurts welfare.

Now let us turn to $\frac{4}{27} \leq c < \frac{1}{4}$. Taxation changes all the tentative cutoffs and we show in Appendix A.3 that the first cutoff x^{1h} decreases and the second cutoff x^{1l} increases in type, while the effect on the third cutoff x^0 is ambiguous. As a result of an increase in c the range of types applying twice shrinks.

We also show that a tax $t > 0$ is always welfare improving in this situation if it changes application behaviour no matter what the observed application pattern is. Intuitively, the positive effect on welfare comes from discouraging workers of types near the first two cutoffs who marginally prefer to send out a second application to do so. From a social perspective the second application is not optimal since the cost is higher than the additional employment created. Increasing cost also has an effect on the third cutoff and although as we just saw for $c > \frac{1}{4}$, discouraging the lowest type workers from applying at all is not beneficial, it turns out that the loss is offset by the gain.

To illustrate the differences in equilibria with high and low cost we consider the following specifications: $M = 1$, $F(x) = x$, $c_1 = 0.2$ and $c_2 = 0.25$ (represented by the red and yellow graph respectively in the figures).

At lower cost c_1 worker types above 0.68 will apply once, workers types in $[0.68, 0.04]$

will apply twice and worker types below 0.04 also apply once. In contrast, at higher cost c_2 all workers apply once. Figure 6 depicts the cutoff types at the intersection of c_1 and the marginal benefit of the second application while c_2 does not intersect with the marginal benefit.

As we would expect workers applying twice with low cost have higher probability of employment compared to with high cost. But the lowest type workers applying once in both cases have lower employment probability given the low cost because the pool of jobs is depleted further by agents of higher type applying twice. This is shown in Figure 7 where employment probabilities decrease continuously in the high cost case but drop sharply when workers switch from two to one application in the low cost case. Figure 8 illustrates that employment ($A(0)$) is lower with high than with low cost.

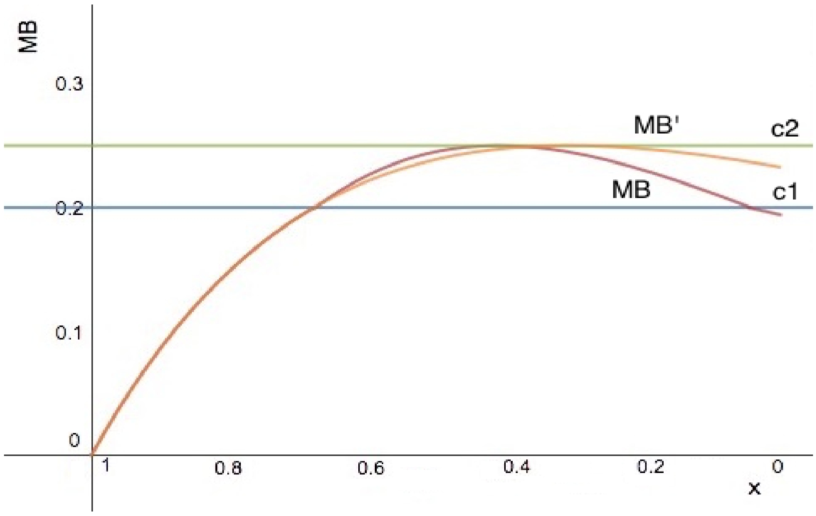

Figure 6: Marginal Benefit of Second Application

We show that for any $c \in (\frac{4}{27}, \frac{1}{4})$, there is $t > 0$ such that welfare strictly increases for economy with $c + t$. There are numerous possible cases. Here we illustrate the proof for the following cases:

1. $k(0) = 0$ and $A(0) = M(1 - c)$
2. $k(0) = 1$, $k(x) = 2$ for some x and $A(0) = 1 - e^{-\frac{1}{M}+t}$ for some t.

Consider 1) $k(0) = 0$ and $A(0) = M(1 - c)$. Total welfare is given by:

$$W = M(1 - c - t) - Mc(\hat{x}^0 + \hat{x}^{1l} - \hat{x}^{1h}) = M(1 - c - t) - Mc(2\hat{x}^{1l})$$

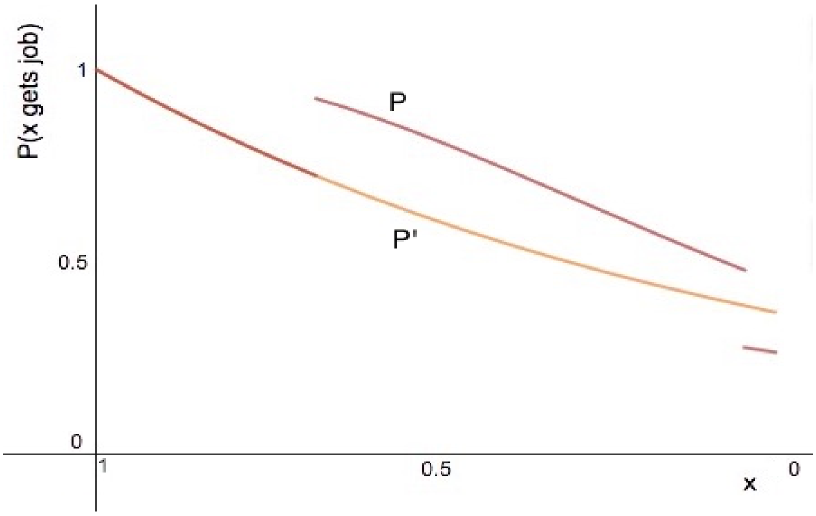

Figure 7: Employment Probability

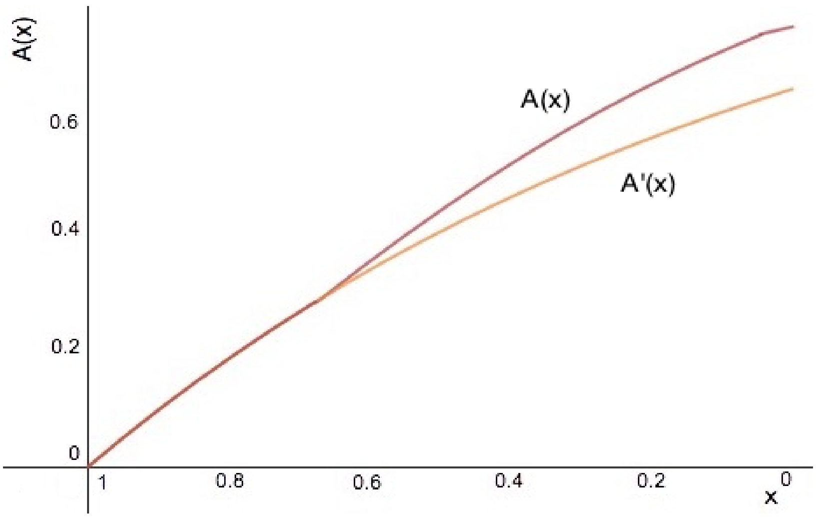

Figure 8: The Accumulation Function

where the first term is total employment and the latter is total cost, consisting of cost of the first application for worker types x, $x < \hat{x}^0$ and the cost of the second application for types in $[\hat{x}^{1h}, \hat{x}^{1l}]$. Taking the derivative with respective to t we obtain:

$$\frac{\partial W}{\partial t} = M\left(-1 - 2c\frac{\partial \hat{x}^{1l}}{\partial t}\right)$$

This is positive if:

$$-2c\frac{\partial \hat{x}^{1l}}{\partial t} > 1$$

$$\frac{1}{2}\frac{(c+t)\left(\frac{3}{2\sqrt{\frac{1}{4}-c-t}}\right) + (2+c+t)}{(c+t)(2+c+t)} > \frac{1}{2c}$$

$$\frac{\left(\frac{3}{2\sqrt{\frac{1}{4}-c-t}}\right)}{(2+c+t)}+\frac{1}{(c+t)}>\frac{1}{c}$$

This is true if evaluated at $t=0$, hence $t>0$ must be welfare improving.

2) $k(0)=1$, $k(x)=2$ for some x and $A(0)=1-e^{-\frac{1}{M}+t}$ for some t implies $\hat{x^0}<\frac{1}{M}$ which is equivalent to $e^{-\frac{1}{M}+t}>c$ and $\hat{x^{1l}}>\frac{1}{M}$. Hence, $e^{-\frac{1}{M}}\in[\beta e^{-\hat{x}^{1l}},e^{-\hat{x}^{1l}}]$. The welfare function and its derivative are given by:

$$W=M(1-e^{-\frac{1}{M}+t})-Mc(1+\hat{x}^{1l}-\hat{x}^{1h})=M(1-e^{-\frac{1}{M}}\delta e^{\hat{x}^{1l}})-Mc(1+\hat{x}^{1l}-\hat{x}^{1h})$$

$$\begin{aligned}\frac{\partial W}{\partial t}&=M\left(-e^{-\frac{1}{M}}e^{\hat{x}^{1l}}\left(\delta'+\delta\frac{\partial\hat{x}^{1l}}{\partial t}\right)-c\left(\frac{\partial\hat{x}^{1l}}{\partial t}-\frac{\partial\hat{x}^{1h}}{\partial t}\right)\right)\\&=M\left(\frac{\partial\hat{x}^{1l}}{\partial t}\left(-e^{-\frac{1}{M}}\delta e^{\hat{x}^{1l}}-c\right)-e^{-\frac{1}{M}}\delta' e^{\hat{x}^{1l}}+c\frac{\partial\hat{x}^{1h}}{\partial t}\right)\\&=M\left(\frac{\partial\hat{x}^{1l}}{\partial t}\left(-e^{-\frac{1}{M}}\delta e^{\hat{x}^{1l}}-c\right)-e^{-\frac{1}{M}}\delta' e^{\hat{x}^{1l}}-c\frac{\beta'}{\beta}\right)\end{aligned}$$

Define $c_t=c+t$ and first note, that the derivative of total employment with respect to tax is always negative:

$$M\left(-e^{-\frac{1}{M}}e^{\hat{x}^{1l}}\left(\delta'+\delta\frac{\partial\hat{x}^{1l}}{\partial t}\right)\right)<0$$

$$\delta'+\delta\frac{\partial\hat{x}^{1l}}{\partial t}>0$$

$$\delta'>\delta\left(\frac{c_t\delta'\frac{3}{2}+\frac{1}{2}(2+c_t)}{(2+c_t)c_t}\right)$$

$$2+c_t>\delta\frac{3}{2}+\frac{\delta\frac{1}{2}(2+c_t)}{\delta' c_t}$$

$$2+c_t>\frac{3}{4}-\frac{3}{2}\sqrt{\frac{1}{4}-c_t}+2\sqrt{\frac{1}{4}-c_t}\frac{(\frac{1}{2}-\sqrt{\frac{1}{4}-c_t})\frac{1}{2}(2+c_t)}{c_t}$$

$$2+c_t>\frac{3}{4}-\frac{3}{2}\sqrt{\frac{1}{4}-c_t}+\left(\frac{1}{2}\sqrt{\frac{1}{4}-c_t}-\left(\frac{1}{4}-c_t\right)\right)\left(\frac{2}{c_t}+1\right)$$

$$2+c_t>\frac{3}{4}-\frac{3}{2}\sqrt{\frac{1}{4}-c_t}+\left(\frac{1}{2}\sqrt{\frac{1}{4}-c_t}-\left(\frac{1}{4}-c_t\right)\right)+\frac{1}{c_t}\sqrt{\frac{1}{4}-c_t}-\left(\frac{1}{2c_t}-2\right)$$

$$0>\frac{1}{2}-\sqrt{\frac{1}{4}-c_t}+\frac{1}{c_t}\sqrt{\frac{1}{4}-c_t}-\frac{1}{2c_t}$$

$$0 > \left(\frac{1}{2} - \sqrt{\frac{1}{4} - c_t}\right)\left(1 - \frac{1}{c_t}\right)$$

Hence for the welfare derivative to be positive for all M it must be positive at its upper bound, i.e.:

$$\frac{\partial W}{\partial t} = \bar{M}\left(-e^{-\hat{x}^{1l}}e^{\hat{x}^{1l}}\left(\delta' + \delta\frac{\partial \hat{x}^{1l}}{\partial t}\right) - c\left(\frac{\partial \hat{x}^{1l}}{\partial t} - \frac{\partial \hat{x}^{1h}}{\partial t}\right)\right) > 0$$

$$\left(\delta' + \delta\frac{\partial \hat{x}^{1l}}{\partial t}\right) - c\left(\frac{\partial \hat{x}^{1l}}{\partial t} - \frac{\partial \hat{x}^{1h}}{\partial t}\right) > 0$$

At $c_t = c$ we have:

$$\left(\delta' + \delta\frac{\partial \hat{x}^{1l}}{\partial t}\right) - c\left(\frac{\partial \hat{x}^{1l}}{\partial t} - \frac{\partial \hat{x}^{1h}}{\partial t}\right) > 0$$

$$\frac{\partial \hat{x}^{1l}}{\partial t}(-\delta - c) - \delta' - c\frac{\beta'}{\beta} > 0$$

$$\frac{1}{2}\frac{c3\delta' + (2+c)}{(c)(2+c)}(\delta + c) - \delta' - c\frac{\beta'}{\beta} > 0$$

$$\delta'\left(\frac{1}{2}\frac{3(\delta + c)}{(2+c)} - 1 + c\frac{1}{\beta}\right) > -\frac{(\delta + c)}{2c}$$

$$\delta'\left(\frac{3\left(\frac{1}{2} - \sqrt{\frac{1}{4} - c} + c\right)}{(2+c)} - 1 - 2\sqrt{\frac{1}{4} - c}\right) > -\frac{(\delta + c)}{c}$$

$$\delta'\left(\frac{3\left(\frac{1}{2} + c\right) - (2+c)}{(2+c)}\right) - \frac{\frac{3}{2} + (2+c)}{(2+c)} > -\frac{\delta}{c} - 1$$

$$\delta'\left(\frac{-\frac{1}{2} + 2c}{(2+c)}\right) - \frac{\frac{3}{2}}{(2+c)} > -\frac{1}{\beta}$$

$$\frac{-\sqrt{\frac{1}{4} - c} - \frac{3}{2}}{(2+c)} > -\frac{1}{\frac{1}{2} + \sqrt{\frac{1}{4} - c}}$$

$$\left(\sqrt{\frac{1}{4} - c} + \frac{3}{2}\right)\left(\frac{1}{2} + \sqrt{\frac{1}{4} - c}\right) < 2 + c$$

$$2\sqrt{\frac{1}{4} - c} + \frac{3}{4} + \frac{1}{4} - c < 2 + c$$

$$\sqrt{1 - 4c} < 1 + 2c$$

This true since $\sqrt{1-4c} < 1 < 1+2c$ and derivative is always positive for the relevant range of $e^{-\frac{1}{M}}$ and $t = 0$.

A.4 Proofs for Section 4

A.4.1 Proof of Proposition 4

Proof. $\gamma < 1$: Define $x' = A_1^{-1}(\gamma A_0(0))$. Note that for some x'

$$\gamma A_0(0) = M_1 \hat{A}\left(\frac{1-0}{M_0}\right) = M_1 \hat{A}\left(\frac{1-x'}{M_1}\right) = A_1(x')$$

Hence $\frac{1}{M_0} = \frac{1-x'}{\gamma M_0}$ by Proposition 2 and $x' > 0$. Since $K_1(x') = K_0(0) > 0$, there exists $\epsilon > 0$ s.t. for workers with types $\check{x} \in [x' - \epsilon, x']$ $K_1(\check{x}) > 0$ and some of them become employed. Hence $A_1(0) > \gamma A_0(0)$.

$\gamma > 1$: Define $x'' = A_0^{-1}(\frac{1}{\gamma} A_1(0))$. Then

$$\frac{1}{\gamma} A_1(0) = M_0 \hat{A}\left(\frac{1-0}{M_1}\right) = M_0 \hat{A}\left(\frac{1-x''}{M_0}\right) = A_0(x'')$$

Hence $\frac{1}{\gamma M_0} = \frac{1-x''}{M_0}$ and $x'' > 0$. Since $K_0(0) > 0$ we have $K_0(x'') > 0$ and all types in $[0, x'']$ make non-zero applications. It is then clear that $\frac{1}{\gamma} A_1(0) = A_0(x'') < A_0(0)$. □

A.4.2 Proof of Proposition 7

Proof. We know from Proposition 5 that with small changes in M, $\gamma < 1$, $\gamma \approx 1$, workers of types between the new and old cutoff types will increase their number of applications by one. For each cutoff type we have $A_1(x_1^j) = \gamma A_0(x_0^j)$. The employment probability of cutoff type $x_1^1 h$ is higher if

$$1 - \left(\frac{A_1(x_1^{1h})}{M_1}\right)^2 > 1 - \frac{A_0(x_1^{1h})}{M_0}.$$

Equivalently,

$$1 - \left(\frac{\gamma A_0(x_0^{1h})}{\gamma M_0}\right)^2 > 1 - \left(\frac{A_0(x_1^{1h})}{M_0}\right)$$

$$\left(\frac{A_0(x_0^{1h}))}{M_0}\right)^2 < \left(\frac{A_0(x_1^{1h})}{M_0}\right)$$

$$\frac{A_0(x_0^{1h})^2}{M_0} < A_0(x_1^{1h})$$

$$\frac{A_0(x_0^{1h})^2}{M_0} < A_0(1 - \gamma(1 - x_0^{1h}))$$

We get the sufficient condition:

$$\frac{1}{1 - x_0^{1h}} \left(1 - A_0^{-1} \left(\frac{A_0(x_0^{1l})^2}{M_0} \right) \right) < \gamma.$$

□