

NBER WORKING PAPER SERIES

EFFICIENCY OF THIN AND THICK MARKETS

Li Gan  
Qi Li

Working Paper 10815  
<http://www.nber.org/papers/w10815>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
September 2004

The views expressed herein are those of the author(s) and not necessarily those of the National Bureau of Economic Research.

©2004 by Li Gan and Qi Li. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Efficiency of Thin and Thick Markets  
Li Gan and Qi Li  
NBER Working Paper No. 10815  
September 2004  
JEL No. J6, J4

**ABSTRACT**

In this paper, we propose a matching model to study the efficiency of thin and thick markets. Our model shows that the probabilities of matches in a thin market are significantly lower than those in a thick market. When applying our results to a job search model, it implies that, if the ratio of job candidates to job openings remains (roughly) a constant, the probability that a person can find a job is higher in a thick market than in a thin market. We apply our matching model to the U.S. academic market for new PhD economists. Consistent with the prediction of our model, a field of specialization with more job openings and more candidates has a higher probability of matching.

Li Gan  
Department of Economics  
University of Texas  
Austin, TX 78712  
and NBER  
gan@eco.utexas.edu

Qi Li  
Department of Economics  
Texas A&M University  
College Station, TX 77843-4228  
qi@econmail.tamu.edu

# Efficiency of Thin and Thick Markets\*

Li Gan<sup>†</sup> and Qi Li<sup>‡</sup>

February 2004

## Abstract

In this paper, we propose a matching model to study the efficiency of thin and thick markets. Our model shows that the probabilities of matches in a thin market are significantly lower than those in a thick market. When applying our results to a job search model, it implies that, if the ratio of job candidates to job openings remains (roughly) a constant, the probability that a person can find a job is higher in a thick market than in a thin market. We apply our matching model to the U.S. academic market for new PhD economists. Consistent with the prediction of our model, a field of specialization with more job openings and more candidates has a higher probability of matching.

Key words: thin and thick market, matching function, market efficiency, empirical test.

## 1 Introduction

In this paper, we are interested in the following question: Compare two markets, one of which has 5 vacancies and 5 applicants, and the other of which has 50 vacancies and 50 candidates. Which market has a lower unemployment rate? The market with a lower unemployment rate is said to be more efficient than the one with a higher unemployment rate.

To answer this question, we set up the following model. We let vacancies have different minimum standards, and job candidates have different productivities. A firm is willing to hire any candidate with a productivity that is higher than its minimum standard, but prefers the candidate with a higher productivity than one with a lower one. A candidate will accept any offer but prefers the firm with a higher minimum standard. Further, we allow the minimum standards and productivities to be randomly drawn from a common distribution, and thereupon to become public information.

---

\*We thank Insik Min, Rich Prisinzano, and Jingyuan Yin for their research assistance. Discussions with Steve Bronars, Dean Corbae, Don Fullerton, Dan Hamermesh, Preston McAfee, Gerald Oettinger, Steve Trejo, and Randy Wright were very helpful. All remaining errors are ours.

<sup>†</sup>Department of Economics, University of Texas, Austin, TX 78712, and NBER. Email: gan@eco.utexas.edu

<sup>‡</sup>Department of Economics, Texas A&M University, College Station, TX 77843-4228. Email: qi@econmail.tamu.edu

For such a market, we prove: (1) The matching probability does not depend on the underlying distribution. (2) When the number of firms and the number of candidates increase, we have a lower unemployment rate (higher matching probability) in the market. For example, when both the number of candidates and of openings are 5, the matching probability is .694. When the number of candidates and openings are 50, the matching probability is .885. (3) The variance of matching probabilities decreases as the market becomes thicker. (4) The conclusion that a thicker market has a larger matching probability than a thinner market continues to hold for the following three more general cases: first, the number of vacancies does not equal the number of candidates; second, the minimum requirement for a vacancy and the productivity of a candidate are random draws from different distributions; and finally, the vacancies and candidates arrive at the market sequentially.

To empirically test our model, we collect information on the U.S. academic market for new PhD economists. We collect the American academic job openings listed for each field in the September, October, November and December issues of *Job Openings for Economists* (JOE), in both 1999 and 2000. We then find out how many of these openings are filled. The ratio of the total number of filled jobs divided by the average openings in each field is the measure of the probability of job matching. In addition, we collect job candidate information from the top 50 departments of economics in U.S. universities.

The empirical estimates strongly support our theoretical hypothesis: a thicker market does have a higher matching probability than a thinner market. In particular, according to the empirical estimate, when the number of candidates and openings are 5, the matching probability is .361. When the number of candidates is 50, the matching probability is .523. Our matching model can also explain the empirical finding of Niederle and Roth (2003) in which a thicker market (through a centralized clearinghouse) increases the matching probability and mobility of gastroenterologists.

Previous literature often uses changes in vacancy rates and in unemployment rates and/or duration to empirically estimate some matching functions.<sup>1</sup> Using the market data for PhD economists offers several advantages over regular job markets. First, there is less of an information problem in this market in the sense that each institute receives applications from almost all potentially qualified job candidates, and almost all job openings are well known to all candidates, as they are published in a single magazine JOE. Second, there is a reasonable consensus in terms of the ranking

---

<sup>1</sup>See, for example, Blanchard and Diamond (1989) and Berman (1997).

of a job, i.e., a job in a better-ranked department is considered by most to be a better job. Third, there is some consensus in terms of the ranking of candidates, although significant heterogeneity still exists.

Although intuitively a thick market has a larger probability of matching, there is no consensus in the literature regarding this intuition. For example, a thicker market has an adverse effect in Burdett, Shi and Wright (2001), has no effect in Lagos (2000), and has a positive effect in Coles and Smith (1998). The different conclusions of these papers result from different matching technology. In Diamond (1982) and Howitt and McAfee (1987), the thickness of a market has externality on itself. If buyers expect that fewer sellers exist in the market, the expected higher cost of transaction discourages buyers from entering into the market. When the same conditions are applied to sellers, low economic activities are expected. Howitt and McAfee (1988) show that it is possible to have multiple externalities (one positive and one negative). These discussions assume that the transaction cost is higher or the probability of matching is smaller in a thinner market. This assumption is intuitively appealing but was neither formally modeled nor empirically verified.

The effect of “thickness” in the market has been studied extensively in the microstructure literature in finance under the term “liquidity.” For example, in Lippman and McCall (1986), a thicker market indicates that more transactions of a homogeneous good take place in a unit of time. In their paper, liquidity is defined in terms of the time elapsed between transactions. This length of time is a function of a number of factors, including the frequency of offers and the flexibility of prices, among others. In an empirical study on common factors that affect liquidity, Chordia, Roll and Subrahmanyam (2000) use five liquidity measures including the difference in prices offered by buyers and sellers, and in quantities offered by buyers and sellers in a period of time. In their approach, the smaller the difference between the prices and the larger the quantities offered by the buyers and sellers of a homogeneous good (an equity), the more liquid a market is. One distinguishing feature in financial markets is that buyers and sellers often arise endogenously. If prices are low, potential sellers easily become buyers. In the labor market, it is hard for workers to become employers or vice versa. Therefore, the pool of employers and employees is often exogenously determined.

Since our model relates the matching probability with the thickness of the market, it can provide a matching function with a microfoundation. The importance of the matching function has been discussed in a recent survey paper by Petrongolo and Pissarides (2001). They claim that both

the matching function and the demand-for-money function are as important as the production function as a tool kit for macroeconomists. However, most of the existing matching functions lack well-received microfoundation as stated in Petrongolo and Pissarides (2001, p.425) “Currently, the most popular functional form, Cobb-Douglas with constant return to scale, is driven by its empirical success and lacks microfoundations. The most popular microeconomic models, such as the urn-ball game, do not perform well empirically.” In this paper, we provide a matching function with microfoundations and show that our matching function performs reasonably well for the empirical data we collected.

This paper is organized as follows: Section 2 introduces our matching model and the basic implications of the model. Section 3 presents our empirical test of the model using the data collected from the U.S. academic job market for new PhD economists. Section 4 concludes the paper.

## 2 The Model

### 2.1 The matching mechanism

Let  $u$  be a measure of the productivity of a job candidate. Let a firm’s profit function be:

$$\pi(u, v) = \max\{0, u - v\}, \tag{1}$$

where  $v$  indicates the type of the firm. In this simple model, the firm will hire a candidate if  $u > v$ , and it prefers a candidate with a higher productivity than one with a lower productivity. The variable  $v$  can be thought of as the firm’s minimum quality requirement.

Without loss of generality we assume that  $v$  takes value in  $[0, C]$  for some  $C > 0$ . We assume a candidate’s utility function is:

$$w(u, v) = \max\{0, v\}. \tag{2}$$

The candidate will accept all job offers as long as  $v \geq 0$  but prefers the firm with higher  $v$  than one with a lower  $v$ .

The matching technology between firms and job candidates defined in (1) and (2) is similar to matching medical interns and residents with hospitals described in Roth (1984). Our primary goal is to examine how the matching probability varies with the number of job candidates and number

of job openings. We consider the problem of  $V$  firms and  $U$  job applicants, and each firm has one opening. Let the firms' minimum requirement be  $v_1, \dots, v_V$ , and job candidates' productivity be  $u_1, \dots, u_U$ . We assume that all productivities and minimum requirements are randomly drawn from a common continuous distribution  $F(\cdot)$  so that no productivity is exactly the same with probability 1 as any minimum requirement. All  $u_i$ 's by all candidates and all firms' minimum requirement  $v_j$ 's are assumed to be known after they are drawn.

According to our model described by (1) and (2), an applicant  $i$  may be hired by firm  $j$  only if the productivity of the applicant  $u_i$  is higher than the minimum requirement of the firm  $v_j$ . If more than one applicant has a higher productivity than the minimum requirement  $v_j$ , firm  $j$  hires the candidate with the highest productivity. Similarly, if more than one firm has lower minimum requirements than applicant  $i$  productivity  $u_i$ , the applicant prefers the firm with the highest minimum requirement. After a match occurs, both the firm and the applicant are out of the market. The process continues until no candidate has a higher productivity than any of the remaining firms' minimum requirements.

In the academic market, a better department is preferred by all newly minted PhDs. The constraint in the market is that each institute has a minimum quality requirement and a better institute has a higher minimum quality requirement. Each institute prefers the candidate with the highest quality that satisfies their minimum requirement. In this case, a necessary condition for a trade to occur between candidate  $i$  and institute  $j$  is  $u_i > v_j$ .

In this matching mechanism, the job applicant who has the highest productivity matches with the firm with the highest minimum requirement, provided that this highest ranked candidate meets the minimum requirement of the firm. Otherwise, the firm leaves the market without filling its vacancy. However, the applicant who does not match with the firm with the highest minimum requirement has additional chances to match with other firms. This matching process repeats in the remaining pool of the applicants and firms.

An alternative way to describe our matching technology is as follows. First we sort all the randomly drawn productivities and minimum requirements. Then a job candidate with the highest productivity matches with the firm with the highest minimum requirement, as long as the minimum requirement is met. Both the job applicant and the firm leave the market. Otherwise, the firm leaves the market. This process is repeated until no firm's minimum requirements can be met by

any remaining candidates.

We assume that firm  $j$  that has a job opening with quality requirement  $v_j$  will be willing to hire any candidate with  $u_i > v_j$ . Let  $v_{(1)} < v_{(2)} < \dots < v_{(V)}$  and  $u_{(1)} < u_{(2)} < \dots < u_{(U)}$  denote the order statistics obtained from  $(v_1, \dots, v_V)$  and  $(u_1, \dots, u_U)$ , respectively. We are interested in the probability that a randomly chosen candidate can find a job. Let  $r$  be the number of people that find jobs,  $0 \leq r \leq n = \min\{V, U\}$ . Let  $\Pr(r)$  denote the probability that exactly  $r$  people find jobs. The average or expected value of  $r$  is given by:

$$M_{U,V} \stackrel{def}{=} E(r) = \sum_{r=0}^n r \Pr(r). \quad (3)$$

In Lemma 1 below, we give the exact probability for a particular order statistic. This lemma is useful for us to obtain matching probabilities.

**Lemma 1:** Let  $u_{(1)} < u_{(2)} < \dots < u_{(U)}$  be the order statistic obtained from i.i.d. data  $u_1, \dots, u_U$ , and  $v_{(1)} < v_{(2)} < \dots < v_{(V)}$  be the order statistic obtained from i.i.d data  $v_1, \dots, v_V$ .  $u_i$  and  $v_j$  have the common distribution  $F(\cdot)$  with pdf  $f(\cdot)$ . Let  $z_{(1)} < z_{(2)} < \dots < z_{(U+V)}$  denote the order statistic obtained from  $(u_{(1)}, \dots, u_{(U)}, v_{(1)}, \dots, v_{(V)})$ . There are  $(U+V)!/[(V!)(U!)]$  such orderings. Let  $\mathcal{Z}_n$  denote the random variable  $(z_{(1)}, \dots, z_{(U+V)})$ . Then for any one particular order  $z$ , we have

$$\Pr(\mathcal{Z}_n = z) = \frac{(V!)(U!)}{(U+V)!}. \quad (4)$$

The proof of Lemma 1 is given in A.1. In Lemma 1, the probability of any particular order statistic  $z$  does not depend on the underlying distribution of candidates and openings. Since the overall matching probability involves accounting the number of appropriate orderings, it does not depend on underlying distributions. In the following sections, we apply Lemma 1 to study how the matching probabilities vary with the number of vacancies and the number of candidates. We first discuss the case where the number of vacancies and the number of candidates are the same, and then we proceed with the case where they are different.

## 2.2 When the number of openings equals the number of candidates

Our primary interest in this paper is to study how the matching probability changes when the number of openings and the number of candidates change. Our discussion starts with the case where the number of applicants is the same as the number of openings. Let  $n = V = U$ , and we



write  $M_{n,n} = M_n$ . We denote by  $A_n$  the probability that a randomly selected person can find a job (when  $V = U$ ), i.e.,

$$A_n = \frac{1}{n} M_n = \frac{1}{n} \sum_{r=1}^n r \Pr(r).$$

We investigate below how  $A_n$  changes as  $n$  changes. We build our model from the simplest case where there is one vacancy and one job applicant.

**The case of  $n = 1$ :**

Let  $u$  and  $v$  be randomly drawn from the same distribution  $f(\cdot)$ . A match occurs if  $u > v$ .

$$A_1 = \int_{\{v < u\}} f(v)f(u)dvdu = \int_{-\infty}^{\infty} \left[ \int_a^u dF(v) \right] dF(u) = \int_{-\infty}^{\infty} F(u)dF(u) = 1/2.$$

In the case of one applicant and one job opening, given that both  $u$  and  $v$  are randomly drawn from the same distribution, the probability that one random draw is larger than the other is  $1/2$ .

**The case of  $n = 2$ :**

Let  $(u_1, u_2)$  be random draws of the two candidates' productivities, and let  $(v_1, v_2)$  be random draws of the minimum requirements of two job openings. All are from the same distribution. Let  $u_{(1)} < u_{(2)}$  be the order statistic of  $(u_1, u_2)$  and  $v_{(1)} < v_{(2)}$  be the order statistic of  $(v_1, v_2)$ . Using Lemma 1 we have:

$$\begin{aligned} \Pr(0) &= \Pr(u_{(1)} < u_{(2)} < v_{(1)} < v_{(2)}) = (2!)^2(1/4!) = 1/6. \\ \Pr(2) &= \Pr(u_{(2)} > v_{(2)}, u_{(1)} > v_{(1)}) \\ &= \Pr(u_{(2)} > v_{(2)} > u_{(1)} > v_{(1)}) + \Pr(u_{(2)} > u_{(1)} > v_{(2)} > v_{(1)}) \\ &= 2\{(2!)^2(1/4!)\} = 1/3. \\ \Pr(1) &= 1 - \Pr(0) - \Pr(2) = 1 - (1/6) - (1/3) = 1/2. \end{aligned}$$

Therefore we have

$$A_2 = \frac{1}{2} \sum_{r=1}^2 r \Pr(r) = [(1/2) + 2(1/3)]/2 = 7/12.$$

We observe that  $A_2 = 7/12 > 1/2 = A_1$ . That is, when the market becomes thicker ( $n$  increases from 1 to 2), the probability that *each* person can find a job is increased from  $1/2$  to  $7/12$ . To understand the intuition of this result, note that since  $\{u_1, u_2\}$  and  $\{v_1, v_2\}$  are from the same

Table 1: Matching Probabilities Based on 100,000 Simulations if ( $V = U$ )  
(Openings and Candidates Have the Same Distributions)

n	1	2	3	4	5	6	7	8	9	10
$A_n$	.5003	.5835	.6337	.6682	.6940	.7135	.7288	.7448	.7562	.7665
Std of $A_n$	.5000	.3434	.2765	.2379	.2116	.1932	.1786	.1661	.1561	.1481
n	20	30	40	50	60	70	80	90	100	1000
$A_n$	.8258	.8543	.8720	.8845	.8938	.9013	.9074	.9124	.9163	.9725
Std of $A_n$	.1041	.0853	.0734	.0657	.0599	.0554	.0523	.0494	.0459	.0144

distribution, the order statistics also have the distribution:  $F_{u_{(1)}}(\cdot) = F_{v_{(1)}}(\cdot)$  and  $F_{u_{(2)}}(\cdot) = F_{v_{(2)}}(\cdot)$ . Given this, we have:

$$\Pr(u_{(1)} > v_{(1)}) = 1/2, \text{ and } \Pr(u_{(2)} > v_{(2)}) = 1/2. \quad (5)$$

If (5) were the only cases that candidates and openings match, we would still end up with a matching probability of 1/2. However, an additional chance exists even when  $u_{(1)} < v_{(1)}$  and  $u_{(2)} < v_{(2)}$  since it is still possible to have  $u_{(2)} > v_{(1)}$ . This additional chance of matching is the source of the effect of a thicker market.

In A.3 and A.4, we calculate the matching probabilities for  $n = 3$  and  $n = 4$ . Although the same approach can be applied to compute  $A_n$  for any  $n > 4$ , the burden of computation becomes tedious as  $n$  increases. A simple alternative is to use simulations to numerically compute  $A_n$ . Let  $A_{n,j}$  be the estimated value of  $A_n$  based on the  $j$ th simulation draw of  $(u_1, \dots, u_n)$  and  $(v_1, \dots, v_n)$ , i.e.,  $A_{n,j}$  equals the number of people finding jobs in the  $j$ th random draw. We estimate  $A_n$  by  $\bar{A}_n = J^{-1} \sum_{j=1}^J A_{n,j}$ . Provided that  $J$  is sufficiently large, we can obtain an estimated value of  $A_n$  with any desired accuracy. We use  $J = 100,000$  in our simulation. We also compute the sample standard error of  $\{A_{n,j}\}_{j=1}^J$  by  $[(J-1)^{-1} \sum_{j=1}^J (A_{n,j} - \bar{A}_n)^2]^{1/2}$ . The results are given in Table 1.

We have already shown that  $A_1 = 0.5$ ,  $A_2 = 7/12 \sim 0.5833$ . In the Appendix we also compute the exact values of  $A_n$  for  $n = 3, 4$ ; they are  $A_3 = 19/30 \sim 0.6333$  and  $A_4 = 187/280 \sim 0.6679$ . Comparing these results with the simulation results of Table 1, we see that the simulation results differ from the theoretical results only in the fourth decimal.

From Table 1 we observe that  $A_n$  increases as  $n$  increases, while the standard error decreases as  $n$  increases. The monotonically increasing relationship between matching probabilities and the thickness of the market can also be clearly seen in Figure 1. The solid line in Figure 1 illustrates the matching probabilities as a function of the number of candidates. As  $n \rightarrow \infty$ , both the candidates and openings become dense in the support of  $f(\cdot)$ . Therefore, the probability of matching is expected to converge to one as  $n \rightarrow \infty$ . This is indeed the case as the next lemma shows.

**Lemma 2:** The employment rate  $A_n$  converges to one as  $n \rightarrow \infty$ .

The proof of lemma 2 is given in the appendix. Note that lemma 2 does not mean that as  $n \rightarrow \infty$ , every individual will find a match. In fact the total unemployed, calculated by  $n(1 - A_n)$ , also goes up as  $n$  increases. For example, when  $n = 10, 100$  and  $1000$ , the average numbers of unemployed workers are roughly 2, 8, and 30, respectively. It is  $1 - A_n$ , the *percentage* of unemployed (the unemployment rate) that goes down as  $n$  increases.

Our theoretical analysis and simulation results show that: (1) A thicker market provides a larger chance of matching; (2) The probability of matching varies less in a thicker market than in a thinner market.

Previous results are obtained by assuming that the job requirements and candidates' productivities have the same distribution. Next, we briefly discuss the case that they have different distributions. We show that in this case the matching probability will depend on the specific distribution functions, but a thicker market still has a larger probability of matching.

We consider the simple case where candidates are randomly drawn from uniform $[0,1]$ , and the openings are randomly drawn from uniform  $[\delta, 1 + \delta]$ ,  $0 \leq \delta \leq 1$ . We will only consider the case of  $V = U = n$ . In Section A.5 we show that:

$$A_1 = \frac{1}{2}(1 - \delta)^2, \text{ and } A_2 = \frac{7}{12}(1 - \delta)^2 + \frac{1}{12}\delta(1 - \delta)^2(2 + 3\delta). \quad (6)$$

Obviously,  $A_2 > A_1$  for all  $\delta \in [0, 1]$ . A thicker market still has a larger probability of matching. For  $n > 2$ , the computation becomes quite tedious. However, one can use simulations to compute  $A_n$  easily for any value of  $n$ . Figure 2 illustrates how the simulated matching probabilities vary with  $n$  and with  $\delta$ . Two patterns emerge from Figure 2. First, as expected, a larger difference in means results in lower matching probabilities. Second, for a fixed value of  $\delta$ , the matching probabilities

increase as the market becomes thicker. By exactly the same argument as in the proof of lemma 2, one can show that as  $n \rightarrow \infty$ ,  $A_n \rightarrow 1 - \delta$  ( $0 < \delta < 1$ ).

Note that when  $\delta \geq 1$ ,  $A_n = 0$  for all  $n$ , because in this case the highest seller's productivity is lower than the lowest buyer's requirement. However, if the two distributions are  $N(\mu, \sigma^2)$  and  $N(\mu + \delta, \sigma^2)$  (the two means also differ by  $\delta$ ), then  $A_n > 0$  for all values of  $\delta$ . This simple example shows that when the two distributions are different, the matching probability will depend on the specific distributions.

### 2.3 The number of vacancies is different from the number of candidates

In the previous section, we only focus on the case where the number of candidates equals the number of openings. In a real market, it is unlikely that there will be exactly the same number of candidates and openings. In this section, we consider cases where the number of candidates is different from the number of openings. They are still random draws from a common distribution.

Let  $U$  be the number of candidates and  $V$  be the number of openings. The number of people who find jobs,  $r$ , must satisfy  $0 \leq r \leq n = \min\{U, V\}$ . Recall that the expected value of  $r$  is

$$M_{U,V} = E(r) = \sum_{r=0}^n r \Pr(r). \quad (7)$$

We summarize some properties of matching functions below.

(i)  $M_{U,V} = M_{V,U}$  is symmetric in  $V$  and  $U$ , (ii)  $M_{U,V}$  increases as either  $V$  or  $U$  increases, (iii) If both  $V$  and  $U$  increase with  $V/U = a$ , a fixed positive constant, then  $B_{U,V} = M_{U,V}/V$  increase as  $V$  ( $U = V/a$ ) increases.

(i) follows from a simple symmetry argument. (ii) is true because adding more candidates or openings to a market obviously cannot reduce the number of matching; in fact, there is a positive probability of increasing the number of matching, thus the average matching of  $M_{U,V}$  will be larger (for any finite values of  $V$  and  $U$ ). (iii) is the most interesting result: it says that when the market becomes thicker, the probability of matching success increases for both candidates and openings. The intuition behind (iii) is quite simple. We have already seen that this is true for the case of  $V = U = n$ . In Appendix A.6 we show how to compute  $M_{U,V}$  (or  $B_{U,V}$ ) for the general  $(U, V)$  case. For example, for  $(U, V) = (1, 2)$  (or  $(2, 1)$ ),  $M_{U,V} = 2/3$ ; for  $(U, V) = (1, 3)$ ,  $M_{U,V} = 3/4$ ; and for  $(U, V) = (2, 4)$ ,  $M_{U,V} = 23/15$ . First we note that

$B_{1,2} = 2/3 < B_{2,4} = (1/2)(23/15) = 23/30$ , so that as the number of  $V$  and  $U$  doubles (the market becomes thicker), the matching probability increases. Next we compare the case of  $(U, V) = (1, 3)$  and  $(2, 2)$ , where we have  $M_{2,2} = 7/6 > 3/4 = M_{1,3}$ . With the *same* total number of openings and candidates, the closer the ratio of  $V/U$  is to 1, the higher the averaging number of people that can find jobs.

Again, a simple alternative is to use simulations to estimate  $M_{U,V}$  ( $B_{U,V}$ ). We will use the simulation method to help us evaluate some of our proposed matching functions in the next section.

## 2.4 A matching function

Because our model relates the matching probability with the thickness of the market, it can provide a matching function with a microfoundation. A series of matching functions has already been introduced in the literature; here we briefly discuss some of the existing matching functions and compare them with our matching function.

In a typical matching model with constant return to scale, the thickness of the market does not enter the matching probability. The relationship between the number of people who are looking for jobs and the number of people who find jobs is different from our claim that market thickness has a positive effect on the job matching ratio. For example, consider a typical matching model with constant return to scale,

$$M = m(U, V) = V \cdot m\left(\frac{U}{V}, 1\right),$$

where  $m(U, V)$  is the matching function,  $M$  is the number of people who find jobs,  $V$  and  $U$  are numbers of job openings and job searchers. The second equality of the previous equation is due to the assumption of the constant return to scale. Rearranging the previous equation, we get:

$$B_{U,V} = \frac{M}{V} = m\left(\frac{U}{V}, 1\right) \tag{8}$$

where  $B_{U,V}$  is firms' matching probability. If the ratio of candidates to vacancies is fixed, so is the matching probability  $M/V$ . A particular form of constant return to scale function is  $M/V = 1 - \exp(-aU/V)$ , which is used in Blanchard and Diamond (1994) where  $a$  is the intensity of the search. Other interesting works related to our matching model include Burdett et al. (2001) and the stock-flow matching of Coles and Smith (1998).

It would be ideal if one could derive an explicit functional form to relate matching probabilities with the thickness of the market. While this goal may be quite difficult to accomplish, we are able to propose a parsimonious approximate matching function which satisfies some basic properties of the theoretical matching function. We will show that this approximate matching function can fit the theoretical matching probabilities very well. We are interested in obtaining a probability matching function, say  $B_{U,V} = M_{U,V}/V$ . However, it is easier to impose restrictions on the matching function  $M_{U,V}$ . We first list some of the properties that  $M_{U,V}$  should preserve.

- (i)  $M_{U,V}$  is symmetric on  $(U, V)$ .
- (ii) For any finite values of  $(U, V)$ ,  $M_{U,V} < \min\{U, V\}$ , and  $M_{U,V}$  is an increasing function in  $U$  ( $V$ ) for a fixed value of  $V$  ( $U$ ).
- (iii) Let  $d = \sqrt{U^2 + V^2}$  denote the distance of  $(U, V)$  to the origin. For  $(U, V) \in \mathbb{R}_+^2$  with  $d = c$ , a constant,  $M_{U,V}$  is monotonically decreasing as  $d(U, V)$  moves away from the middle point of  $V = U$  (along the arc of  $d = c$ ).

The following simple matching function satisfies the above four conditions:

$$M_{U,V}^{(0)} = \alpha_0 + \alpha_1 \min\{U, V\} + \frac{\alpha_2}{d}, \quad (9)$$

where  $d = \sqrt{V^2 + U^2}$ , and  $\alpha_0, \alpha_1$ , and  $\alpha_2$  are parameters ( $\alpha_1$  is positive and  $\alpha_2$  is negative).

It is obvious that  $M_{U,V}^{(0)}$  in (9) satisfies properties (i) and (ii) above. To see that it also satisfies (iii), note that when  $d = c$  is a constant,  $M_{U,V}^{(0)} = \alpha_1 \min\{U, V\}|_{d=c} + \alpha_2/c$ , which decreases monotonically as  $(U, V)$  moves away from the middle point of  $U = V$  (along the arc of  $d = c$ ).

By rearranging (9) in terms of matching probability, we get

$$\frac{M_{U,V}^{(0)}}{V} = \alpha_0 + \alpha_1 \min\left\{\frac{U}{V}, 1\right\} + \frac{\alpha_2}{Vd}. \quad (10)$$

If we replace  $\alpha_2/(Vd)$  by  $\alpha_2/d$  (removing the  $1/V$  factor) in (10) we obtain the following alternative approximate matching function:

$$\frac{M_{U,V}^{(1)}}{V} = \alpha_0 + \alpha_1 \min\left\{\frac{U}{V}, 1\right\} + \frac{\alpha_2}{d}. \quad (11)$$

Table 2: Regressions of Matching Probabilities

Models	$1 \leq U, V \leq 10$		
	Model (10)	Model (11)	Model (12)
constant	.015 (.841)*	.075 (4.28)	.060 (1.96)
$\min\{U/V, 1\}$	.805 (37.7)	.787 (42.8)	.779 (42.3)
$1/(U^2 + V^2)^{1/2}$		-.423 (-7.07)	-.456 (-4.55)
$1/[V(U^2 + V^2)^{1/2}]$	-.281 (-3.20)		
$\alpha_3$			1.19 (3.05)
$R^2$	.937	.951	.957
Number of observations	100	100	100

\* t-values are in parentheses.

Interestingly we observe that (11) performs better than model (10) using both theoretical (simulated) matching probabilities and the empirical data.

A more flexible model than (11) is

$$\frac{M_{U,V}^{(1)}}{V} = \alpha_0 + \alpha_1 \min\left\{\frac{U}{V}, 1\right\} + \frac{\alpha_2}{d^{\alpha_3}}. \quad (12)$$

When  $\alpha_3 = 1$ , (12) reduces back to (11). Because model (12) is nonlinear in parameters, one needs some iterative procedure such as a nonlinear least squares method to estimate model (12).

To examine how well our proposed matching functions approximate the theoretical (simulated) matching function, we carry out a least squares regression, using (simulated) theoretical values of  $B_{U,V} = M_{U,V}/V$  as the dependent variable, and estimate models (10), (11), and (12). The regression results for models in (10) to (12) are reported in Table 2.

As can be seen in Table 2, our specifications can explain the (simulated) theoretical matching probability well, with  $R^2$  being at least .937. The  $R^2$  is .951 for model (11) and .957 for model (12). The results show that all of the proposed models fit the theoretical model very well. However, this does not imply that one should expect that all of them should fit empirical data equally well. As we will see shortly, model (12) provides the best fit for the empirical data we collected.

## 2.5 A sequential matching mechanism

Up to now we have only considered a static model where all candidates and openings arrive at the market simultaneously. In this section we briefly discuss the case of a sequential matching model. Our approach follows closely that of Coles and Smith (1998) who capture a realistic feature of market search, that if a job seeker cannot match with the exiting pool of vacancies, he/she will wait for the arrivals of new job vacancies. We consider two extreme cases: (i) All matched pairs can break up an earlier match and re-match in a later period without a cost. The time discount rate is zero. (ii) Both the re-match cost and the time discount rate are infinite.

### **A zero re-matching cost and a zero time discount rate**

It is easy to see that in this case the results of sections 2.2 and 2.3 remain valid without changes. Suppose at period  $t$ , we have a cumulative of  $U_t$  job candidates, and a cumulative of  $V_t$  vacancies, the number of matches will be exactly the same as in the static case with a total number of  $U_t$  candidates and  $V_t$  vacancies. This is because all matched pairs can freely break up with earlier matches and to find the best match available to them. The highest quality individual will match with the best job available provided her quality meets the minimum requirement of that job. The second highest quality individual will match with the next best available job. Consequently, the matching results will be identical as in the static case with the same total numbers of candidates and vacancies.

### **An infinite re-matching cost and an infinite time discount rate**

When both the cost of re-entering the market and the time discount rate are infinite, all firms and individuals will try to find a match as soon as possible, and when a match is found, the matched pair will exit the market. Although these assumptions are not realistic, they serve as a benchmark case and from which we can deduct useful information on the more realistic finite re-matching cost/discount factor cases.

We will only consider the case where the number of candidates equals the number of openings. This will be the case if candidates and openings arrive at the market in pairs so that the total numbers of candidates and openings equal each other at all times. We assume that different pairs arrive at the market sequentially. If the first pair of candidate and opening matches with each other, they will sign a contract and exit the market. If not, they become stock and wait for



matching opportunities among future arrivals. When a pair (a candidate and an opening) arrives at the market, the two can match with each other or match with the existing stock, according to whichever gives the higher utility. If no match is found, they become stock.

Let  $n = V = U$  denote the total number of openings and candidates. We use  $\bar{P}(r)$  to denote the probability that exactly  $r$  people find jobs, and use  $\bar{A}_n = n^{-1} \sum_{r=0}^n r \bar{P}(r)$  to denote the mean value of  $r$  (the bar notation is used to emphasize a sequential matching process).

**The case of  $V = U = 1$ .**

In this case we have  $\bar{P}[U > V] = 1/2$  as before, which gives  $\bar{A}_1 = \bar{P}(1) = 1/2$ .

**The case of  $V = U = 2$ .**

Let  $u_j$  ( $v_j$ ) be the  $j$ th arrival of candidates (openings),  $j = 1, 2$ . Because candidates and openings arrive sequentially, we cannot use the order statistics to compute  $\bar{P}(r)$ . However, the result obtained earlier can help calculate the matching probabilities.

The total number of different rankings of  $u_1, u_2, v_1$  and  $v_2$  is  $4! = 24$  (four times that of the order statistic case). We can use the calculation of  $\Pr(r)$  to help us to obtain  $\bar{P}(r)$ . For example, in the market of simultaneous arrival, the order statistic that no one finds a job is:  $u_{(1)} < u_{(2)} < v_{(1)} < v_{(2)}$ , and the probability is  $\Pr(0) = 1/6$ . In the market of sequential arrival, there are four cases that no one finds a job: (i)  $u_1 < u_2 < v_1 < v_2$ , (ii)  $u_2 < u_1 < v_1 < v_2$ , (iii)  $u_1 < u_2 < v_2 < v_1$ , and (iv)  $u_2 < u_1 < v_2 < v_1$ , giving  $\bar{P}(0) = 4/24 = 1/6$ . So the probability that no one finds a job remains unchanged.

There is only one case that results in different matching probabilities between a sequential market and a simultaneous market. In the case of  $v_1 < u_2 < v_2 < u_1$ ,  $u_1$  will match with  $v_1$  and then  $(u_1, v_1)$  exit the market. In the second period,  $v_2$  and  $u_2$  arrive at the market but they cannot match because  $u_2 < v_2$ . If the two pairs had arrived simultaneously, there would be two matched pairs,  $u_1$  with  $v_2$ , and  $u_2$  with  $v_2$ . So we see that when arrivals are sequential, the matching probability decreases and the market becomes less efficient. Using (4) we obtain:

$$\bar{P}(0) = 4/24 = 1/6 \quad (= 4/24 \text{ as in the simultaneous arrival case})$$

$$\bar{P}(1) = (12 + 1)/24 = 13/24 \quad (\text{it was } 12/24 = 1/2 \text{ in the simultaneous arrival case})$$

$$\bar{P}(2) = (8 - 1)/24 = 7/24 \quad (\text{it was } 8/24 = 1/3 \text{ in the simultaneous arrival case})$$

$$\text{Thus, } \bar{A}_2 = (1/2) \sum_{r=0}^2 r \bar{P}(r) = [(13/24) + 2(7/24)]/2 = 27/48 > 1/2 = \bar{A}_1.$$

We still observe that as the total number of candidates and openings goes up, the average matching probability increases. However,  $\bar{A}_2 = 27/48 < 28/48 = A_2$ , the market of sequential arrival is less efficient compared to the case that all the candidates and openings arrive simultaneously, which is an expected result since sequential trading may lead to a very high quality candidate to match with a vacancy with a very low minimum requirement, resulting in a less efficient market.

Let  $\Pr[(u_i, v_j)]$  denote the probability that  $u_i$  matches  $v_j$ . Then conditional on  $u_1 < v_1$  (so that  $u_1$  and  $v_1$  become stock), it is easy to show that  $\Pr[(u_2, v_2)] = 1/3 > \Pr[(u_2, v_1)] = 1/4$ . Thus, our matching mechanism implies that  $u_2$  has a lower matching probability, or higher rejection rate, when meeting with  $v_1$  (an opening from the stock) than when meeting with  $v_2$  (a random draw from the distribution of job openings). This is because the openings from the stock have a higher mean value than those drawn from the population. As is easily shown, our model implies that the rejection rate between a candidate from the flow and an opening from the stock increases as the size of the stock increases, or equivalently as the averaging matching probability increases (as argued by Petrongolo and Pissarides (2001, p.406)). This is because the mean of the stock of openings goes up as its size goes up, resulting in a higher rejection rate for a given pair. Nevertheless, the average matching probability still goes up since there are more matching opportunities as the market gets thicker.

Table 3 reports the simulated values of  $\bar{A}_n$  for  $n$  from 1 to 1,000 (based on 100,000 replications). Since  $\bar{A}_2 = 27/48 = .5625$ , we see again that the simulated value matches the true value in the first three decimals.

In Table 3, we observe similar phenomena in the case of simultaneous trading, i.e.,  $\bar{A}_n$  increases (with a decreasing rate) while the standard deviation of  $\bar{A}_n$  decreases as  $n$  increases. The dashed line in Figure 1 shows that, as expected, the  $\bar{A}_n$  curve is lower than the solid line of the  $A_n$  curve. A larger friction exists in a market of sequential arrivals.

Even though sequential arrival leads to higher friction, the intuition that a thick market is more efficient than a thin market remains the same as in the case of a simultaneous arrival. Further it can be shown that  $\bar{A}_n \rightarrow 1$  as  $n \rightarrow \infty$ .

So far we consider two extreme case: zero time discount rate and re-match cost vs. infinite time discount rate and re-match cost. Complete discussions of more realistic cases where the time discount rate and the re-match cost are some finite positive numbers are left for future research.

Table 3: Matching Probabilities Based on 100,000 Simulations if ( $V = U$ )  
(Vacancies and Candidates Arrive Sequentially)

n	1	2	3	4	5	6	7	8	9	10
$\bar{A}_n$	.5003	.5622	.5999	.6257	.6452	.6603	.6728	.6833	.6922	.7001
Std of $\bar{A}_n$	.5001	.3330	.2665	.2273	.2034	.1846	.1694	.1575	.1475	.1403
n	20	30	40	50	60	70	80	90	100	1000
$\bar{A}_n$	.7486	.7736	.7905	.8030	.8127	.8207	.8272	.8332	.8383	.9248
Std of $\bar{A}_n$	.0977	.0812	.0713	.0632	.0582	.0538	.0515	.0486	.0461	.0146

We did not yet consider searching cost in our model. Adding a fixed searching cost will not alter any of the conclusions obtained earlier. A variable searching cost may reduce matching probability. Given the rapid improvement of internet searching, it seems that a fixed search cost is appropriate for most situations such as economics new Ph.D's market. We leave more detailed discussions on extensions such as strategic trading behavior and variable searching cost to future research work.

### 3 Data Collection and an Empirical Test

Empirical study of this issue can be very difficult. It is relatively simple to collect information about successfully completed transactions in a particular market. However, gathering data about the participants who failed to complete transactions is often very difficult.

The job market for new PhD economists, therefore, provides an excellent opportunity for exactly such an empirical study of thin and thick markets' performances. First, we must identify the levels of supply and demand for this market. The information we require to determine market demand is available through the journal *Job Openings for Economists* (JOE). The problem of information asymmetry is minimal when we consider the job market in economics because JOE provides virtually complete information sets for the supply of the academic jobs in the U.S. In other labor markets, we often do not know what specific information sets job applicants can access; but in this case we do because the journal is widely available to candidates going on the job market. In addition, we may determine the level of market supply by contacting graduate programs in economics regarding

their PhDs who have gone on the job market in the past several years.

Our data is organized by field. The definition of the field can be found in the “Classification System of Journal Articles” by the *Journal of Economic Literature*. In particular, we use the field consisting of a capital letter and a numeral. For example, E0 means “Macroeconomics and Monetary Economics.”

We collect the American academic job openings listed in the September, October, November, and December issues of JOE in 1999 and 2000. In JOE,  $m$  job openings are listed with  $n$  fields where  $m$  and  $n$  are integers. We determine that each field has average openings of  $m/n$ . For example, in 2000, the Department of Economics at the Texas A&M University had five openings in nine different fields. We assign each field with  $5/9$  openings. We sum all the average openings for all American universities by each field. We then find out how many of these openings are filled by going to each institute’s website and/or by contacting relevant people. The ratio of the total number of filled jobs divided by the average openings in each field is the measure of the probability of job matching.

In addition, we collect the job candidate information. We search the links of job candidates in each of the top-50 departments in the U.S. defined in Dusansky and Vernon (1998). We use the first field listed in each candidate’s CV or in the brief introduction of a candidate if no CV is available.

The summary information of the markets is listed in Table 4. In addition to showing the ten fields with the most job openings in the table, we include any field (AF), the mean of the remaining fields not listed in the table, and the whole market. In both years, AF is by far the largest “field.” Macroeconomics (E0), Microeconomics (D0), and International Economics (F0) were the top three fields other than AF in both years. In 1999, the mean of the matching probabilities in the ten fields with the most job openings is .501, while the mean of the matching probabilities in the remaining fields is .305. Thicker fields do have larger matching probabilities than thinner fields. The same pattern repeats in 2000 where the mean of the matching probabilities for the ten fields with the largest demand is .451, while the rest of the fields have an average matching probability of .268.

We estimate our proposed matching models, given in (10), (11), and (12), using the collected data. We are primarily interested in the sign of the coefficient for the variable of thickness, measured by the the variable  $d = (\text{candidates}^2 + \text{openings}^2)^{1/2}$ .

Table 4: Summary of Academic Markets for New PhD Economists

Fields with most openings	average openings	# of filled positions	probability of matching	number of candidates
<u>Year 2000</u>				
Any field(AF)	95.3	42	.441	0
Macro(E0)	49.3	30	.608	83
Micro(D0)	36.2	18	.498	42
International(F0)	34.9	25	.717	39
Econometrics(C1)	33.5	13	.388	43
Financial Econ(G0)	33.4	19	.568	39
Agric Econ(Q0)	25.6	13	.507	16
Public Econ(H0)	25.4	9	.354	37
General Econ(A1)	22.6	4	.177	0
Health Econ(I1)	21.2	9	.426	11
IO(L0)	19.6	15	.765	60
Mean of remaining fields	2.96	1.40	.305	2.16
Total	617	308	.499	529
<u>Year 2001</u>				
Any field(AF)	125.0	64	.512	0
Macro(E0)	54.8	31	.566	72
International(F0)	39.6	11	.277	29
Micro(D0)	38.2	20	.523	34
Agric Econ(Q0)	37.9	13	.343	11
Econometrics(C1)	36.0	13	.361	32
Health Econ(I1)	34.5	18	.521	14
Financial Econ(G0)	31.9	16	.501	24
Public Econ(H0)	20.7	11	.532	25
IO(L0)	20.7	15	.726	59
General Econ(A1)	18.3	3	.164	0
Mean of remaining fields	3.27	1.20	.268	1.92
Total	696	308	.443	445

Table 5 gives the estimation results of models (10), (11), and (12), in the same format as that of Table 2. It is clear that the regressions based on model (12) have the best fit, followed by model (11). In all these different specifications using different sample data, the parameter estimates of  $\alpha_2$ , the coefficient of the inverse of the thickness variable  $d$ , are negative. Moreover, they are significant at the 5% level for eight out of nine cases, and are all significant at the 10% level (note that it is an one-sided test). Thus our estimation results predict that the matching probability increases as the market becomes thicker, consistent with the main prediction of our theoretical model. In other words, a thicker market produces a better probability of matching. The statistically insignificant estimates of  $\alpha_1$  reflects the fact that the total number of supply in each field,  $U$ , is measured with error.<sup>2</sup>

Figure 3a gives the estimated curves using 2000 job market data, Figure 3b uses 2001 job market data, and Figure 3c uses the pooled sample. These figures graph the observed and predicted matching probabilities for models (11) and (12). Each point in these figures represents one field. The dotted line in each figure represents the predicted probability based on model (11); the solid line plots the predicted probability based on model (12). The prediction is carried at the sample mean of  $\min\{U/V, 1\}$ . Comparing model (11) with model (12) we observe that the nonlinear model shows more pronounced thickness effects. Within the same model (say model (12)), all three graphs are similar, reflecting the fact that estimates from different sample are similar. From all the graphs, we clearly see that matching probability is an increasing and concave function of the thickness ( $d$ ) of the market.

To understand the magnitude of the effect of thickness on the matching probability, consider model (12) where the number of candidates equals the number of job openings; we have that (when  $U = V$ ) the matching probability is

$$(\widehat{M/V}) = \hat{\alpha}_0 + \hat{\alpha}_1(1) + \hat{\alpha}_2/d^{\hat{\alpha}_3},$$

where  $d = \sqrt{2V}$  (since  $U = V$ ). For  $U = V = 5, 10, \text{ and } 50$ , and using the pooled sample estimation

---

<sup>2</sup>The total number of candidates in these 50 schools in each field is our noisy measure of the total supply of the market. One source of noise comes from the classification of candidates' fields. For example, students often indicate their fields to be one of the thicker fields. In the two year period we have data, among the 126 fields that have academic openings, only 45% of them have candidates, although 62% of those fields have some success to hire at least one candidate.

result, our empirical model predicts the matching probabilities of .361, .421, and .523, respectively.

Finally, in order to check whether fields such as “any field” (AF) and “general economics” (A1) contaminate our estimation results, we conduct estimates removing fields “AF” and “A1”. These two fields have large numbers of openings while there are no candidates labeled as “any field” or “general economics.” Estimation results not reported here (they are available upon request) show virtually identical parameters estimates as well as the goodness-of-fit  $R^2$  for the results given in Table 5. Thus the fact that there are zero candidates in the thick fields “AF” and “A1” does not affect our estimation results nor the conclusions derived from them.

## 4 Conclusions

In this paper we propose a matching model with the matching probability depending on the thickness of a market. In our model, the minimum job requirements of firms and the productivities of job applicants are randomly drawn from a common distribution. A firm employs a job applicant only if the job applicant’s productivity is higher than the minimum requirement. All firms prefer a higher productivity applicant to a lower one, and all applicants prefer a higher minimum standard firm to a lower one. In this hypothetical market, we show that the probabilities of matches in a thin market differ significantly from those in a thick market.

We also characterize the case where firms and applicants have different distributions, the case where the number of openings does not equal the number of applicants, and the case where openings and candidates arrive at the market sequentially. In all these cases, the matching probability still increases with the thickness of the market. In addition, we propose a parsimonious matching function which is fairly close to the (simulated) theoretical matching function.

The implications of our model are consistent with the liquidity literature in the financial market where more trading occurs in a thicker market than in a thinner market. We apply our matching model to the U.S. academic market for junior PhD economists. Consistent with the prediction of our model, a field with more job openings and more candidates has a higher probability of matching. In particular, according to our model, the matching probability increases from .361 for 5 candidates and openings to .523 for 50 candidates and openings.

The model above can be extended in many directions, such as to the regular labor or housing

Table 5: Regressions of Matching Probabilities  
(U.S. Academic Market for New PhD Economists)

	Model (10)	Model (11)	Model (12)
<u>Job Market in January 2000</u>			
constant	.284 (5.69)*	.362 (6.67)	.847 (1.98)
min{U/V, 1}	.208 (2.51)	.142 (1.75)	.020 (.237)
$1/[(V^2 + U^2)^{1/2}]$		-.112 (-4.03)	-.674 (-1.52)
$1/[V(V^2 + U^2)^{1/2}]$	-.017 (-2.72)		
$\alpha_3$			.223 (1.30)
$R^2$	.250	.339	.466
# of observations	61	61	61
<u>Job Market in January 2001</u>			
constant	.370 (8.29)	.456 (8.97)	.658 (3.36)
min{U/V, 1}	-.023 (-.29)	-.094 (-1.19)	-.155 (-1.84)
$1/[(V^2 + U^2)^{1/2}]$		-.212 (-4.78)	-.468 (-2.20)
$1/[V(V^2 + U^2)^{1/2}]$	-.061 (-3.80)		
$\alpha_3$			.411 (1.70)
$R^2$	.192	.273	.315
# of observations	65	65	65
<u>Pooled Sample of 2000 and 2001</u>			
constant	.314 (9.43)	.398 (10.88)	.775 (3.36)
min{U/V, 1}	.106 (1.82)	.035 (.62)	-.068 (-1.15)
$1/[(V^2 + U^2)^{1/2}]$		-.139 (-5.97)	-.591 (-2.45)
$1/[V(V^2 + U^2)^{1/2}]$	-.023 (-3.94)		
$\alpha_3$			.274 (2.03)
$R^2$	.160	.266	.369
# of observations	126	126	126

\* t-values are in parentheses.



markets. It may require more serious effort to relax some assumptions made in this paper. For example, job candidates may have different preference orderings of employers, as may employers. More general transferable utility functions may be more appropriate for a general labor market because one may enjoy a higher utility if he/she moves to a better quality firm. Buyers in the housing market have several aspects to consider, while sellers may only care about the selling prices.

## A Appendix

### A.1 Proof of Lemma 1

*Proof:* The joint distribution of  $(u_{(1)}, \dots, u_{(U)})$  and  $(v_{(1)}, \dots, v_{(V)})$  are  $(U!) \prod_{r=1}^U f(u_{(r)})$  and  $(V!) \prod_{r=1}^V f(v_{(r)})$ , respectively. Therefore, the joint distribution function for  $(u_{(1)}, \dots, u_{(U)}, v_{(1)}, \dots, v_{(V)})$  is  $(V!)(U!)[\prod_{r=1}^V f(u_{(r)})][\prod_{r=1}^U f(v_{(r)})]$ .

$$\begin{aligned}
\Pr(\mathcal{Z}_n = z) &= P(z_{(1)} < z_{(2)} < \dots < z_{(U+V)}) = (V!)(U!) \int_{\{z_{(1)} < z_{(2)} < \dots < z_{(U+V)}\}} \prod_{r=1}^{U+V} dF(z_{(r)}) \\
&= (V!)(U!) \int_{\{z_{(2)} < z_{(3)} < \dots < z_{(U+V)}\}} F(z_{(2)}) \prod_{r=2}^{U+V} dF(z_{(r)}) \\
&= (V!)(U!) \int_{\{z_{(3)} < \dots < z_{(U+V)}\}} (1/2)F^2(z_{(3)}) \prod_{r=3}^{U+V} dF(z_{(r)}) \\
&= (V!)(U!) \int_{\{z_{(3)} < \dots < z_{(U+V)}\}} (1/2) \dots (1/(U+V-1)) \int_a^b F^{U+V-1}(z_{(U+V)}) F(z_{(U+V)}) \\
&= \frac{(V!)(U!)}{(U+V)!}.
\end{aligned}$$

### A.2 Proof of Lemma 2

We know that  $A_n$  is independent of the distribution  $f$ . Therefore, without loss of generality we assume that  $f$  is a uniform distribution in the unit interval. For any (small)  $\eta > 0$ , we choose  $m = \lceil 1/\eta \rceil > 1/\eta$ . ( $\lceil \cdot \rceil$  denotes the integer part of  $\cdot$ ) and divide the unit interval into  $m$  intervals with equal length  $1/m$  for each. That is:  $[0, 1] = \cup_{l=1}^m I_l$ , where  $I_l = [(l-1)/m, l/m)$  ( $l = 1, \dots, m$ , with  $I_m = [(m-1)/m, 1]$ ). Let  $n_{u,l}$  and  $n_{v,l}$  denote the number of observations from  $\{u_i\}_{i=1}^n$  and  $\{v_i\}_{i=1}^n$  that fall inside in interval  $I_l$  ( $l = 1, \dots, m$ ). We know that on the average there are  $n/m$  observations from both  $\{u_i\}_{i=1}^n$  and  $\{v_i\}_{i=1}^n$  that fall inside interval  $I_l$  for all  $l = 1, \dots, m$ . In fact by the strong law of large number (Billingsley 1986(p.80)) we have  $P(\lim_{n \rightarrow \infty} n_{s,l}/n = 1/m) = 1$  for all  $l = 1, \dots, m$  ( $s = u, v$ ).

Note that the candidates with  $u_i$ 's fall inside the interval  $I_l$  can match with any job opening with  $v_j$ 's falls in  $I_{l-1}$  ( $l = 2, \dots, m$ ). Given that with probability one that  $n_{u,l}/n \rightarrow 1/m$  and  $n_{v,l-1}/n \rightarrow 1/m$ , we know that, with probability approaching to one as  $n \rightarrow \infty$ , that there can have  $n/m$  matches for  $u_i$ 's  $\in I_l$  matching with  $v_j$ 's  $\in I_{l-1}$ . Sum over  $l$  from 2 to  $m$  we get, with probability one, that the number of matched candidates is at least (since we ignore the possibility that  $u_i$ 's  $\in I_1$  may also find match)  $n_{match}/n \geq (m-1)/m \geq 1-\eta$ , or more formally, we have, as  $n \rightarrow \infty$ ,

$$P\left(\frac{n_{match}}{n} \geq 1-\eta\right) \rightarrow 1. \quad (13)$$

Therefore we have

$$1 \geq A_n = \frac{1}{n} \sum_{r=0}^n rP(r) \geq \frac{1}{n} \sum_{r \geq n(1-\eta)} rP(r) \rightarrow 1,$$

because for any  $1 > \epsilon > 0$ , we can choose  $\eta = \epsilon/2$  and by (13), we have  $n^{-1} \sum_{r \geq n(1-\eta)} rP(r) \geq n^{-1}n(1-\eta) \sum_{r > n(1-\eta)} P(r) \geq (1-\eta)^2 \geq 1-\epsilon$ . Thus,  $n^{-1} \sum_{r \geq n(1-\eta)} rP(r) \rightarrow 1$  as  $n \rightarrow \infty$  which implies  $A_n \rightarrow 1$ , completing the proof of Lemma 2.

Let  $\Pr(\# \geq r)$  denote the probability that at least  $r$  people find jobs. Then it is easy to see that  $\Pr(\# \geq r) = \Pr(u_{(n)} > v_{(n-r+1)}, u_{(n-1)} > v_{(n-r)}, \dots, u_{(n-r)} > v_{(2)}, u_{(n-r+1)} > v_{(1)})$ . The next lemma shows that  $\Pr(\# \geq r)$  can be used to compute  $E(r)$ .

**Lemma 3:** Let  $\#$  denote the number of people who find jobs ( $0 \leq \# \leq \min\{V, U\}$ ), and denote by  $\Pr(\# \geq r) = \sum_{m=r}^n \Pr(m)$  the probability that at least  $r$  people find jobs. Then

$$E(r) = \sum_{r=1}^n \Pr(\# \geq r).$$

*Proof:*

$$\begin{aligned} E(r) &= \sum_{r=1}^n r \Pr(r) = \{\Pr(1) + (2) \Pr(2) + \dots + (n) \Pr(n)\} \\ &= \{[\Pr(1) + \Pr(2) + \dots + \Pr(n)] + [\Pr(2) + \dots + \Pr(n)] + [\Pr(n-1) + \Pr(n)] + \Pr(n)\} \\ &= \sum_{r=1}^n \Pr(\# \geq r). \end{aligned}$$

### A.3 The case of $n = 3$

Let  $u_3 > u_2 > u_1$  be the order statistic of candidates, and  $v_3 > v_2 > v_1$  be the order statistic of openings (we omit the parentheses in the subscripts to simplify the notation).

$$\Pr(0) = \Pr(u_3 < u_2 < u_1 < v_1 < v_2 < v_3) = (3!)^2(1/6!) = 1/20.$$

$$\Pr(\# \geq 1) = 1 - \Pr(0) = 19/20.$$

$$\Pr(\# \geq 2) = \Pr(u_3 > v_2, u_2 > v_1) = \Pr(u_3 > v_2 > u_2 > v_1) + \Pr(u_3 > u_2 > v_2 > v_1) = 14/20$$

since

$$\begin{aligned}
\Pr(u_3 > v_2 > u_2 > v_1) &= \Pr(u_3 > v_3 > v_2 > u_2 > v_1 > u_1) + \Pr(v_3 > u_3 > v_2 > u_2 > v_1 > u_1) \\
&+ \Pr(u_3 > v_3 > v_2 > u_2 > u_1 > v_1) + \Pr(v_3 > u_3 > v_2 > u_2 > u_1 > v_1) \\
&= 4[(3!)/(6!)] = 4/20.
\end{aligned}$$

$$\begin{aligned}
\Pr(u_3 > u_2 > v_2 > v_1) &= \Pr(u_3 > u_2 > v_3 > v_2 > v_1 > u_1) + \Pr(u_3 > v_3 > u_2 > v_2 > v_1 > u_1) \\
&+ \Pr(v_3 > u_3 > u_2 > v_2 > v_1 > u_1) + \Pr(u_3 > u_2 > v_3 > v_2 > u_1 > v_1) \\
&+ \Pr(u_3 > v_3 > u_2 > v_2 > u_1 > v_1) + \Pr(v_3 > u_3 > u_2 > v_2 > u_1 > v_1) \\
&+ \Pr(u_3 > u_2 > v_3 > u_1 > v_2 > v_1) + \Pr(u_3 > v_3 > u_2 > u_1 > v_2 > v_1) \\
&+ \Pr(v_3 > u_3 > u_2 > u_1 > v_2 > v_1) + \Pr(u_3 > u_2 > u_1 > v_3 > v_2 > v_1) \\
&= 10[(3!)/(6!)] = 10/20.
\end{aligned}$$

$$\begin{aligned}
\Pr(3) &= \Pr(u_3 > v_3, u_2 > v_2, u_1 > v_1) \\
&= \Pr(u_3 > v_3 > u_2 > v_2 > u_1 > v_1) + \Pr(u_3 > u_2 > v_3 > v_2 > u_1 > v_1) \\
&+ \Pr(u_3 > v_3 > u_2 > u_1 > v_2 > v_1) + \Pr(u_3 > u_2 > v_3 > u_1 > v_2 > v_1) \\
&+ \Pr(u_3 > u_2 > u_1 > v_3 > v_2 > v_1) \\
&= 5\{(3!)^2/(1/6!)\} = 5/20.
\end{aligned}$$

Therefore, by Lemma 3 we have

$$A_3 = \frac{1}{3} \sum_{r=1}^3 \Pr(\# \geq r) = [(19/20) + (14/20) + (5/20)]/3 = 19/30.$$

#### A.4 The case of $n = 4$

$$\Pr(0) = (4!)^2/(8!) = 1/70 \text{ by Lemma 1. } \Pr(\# \geq 1) = 1 - P(0) = 69/70.$$

$$\Pr(\# \geq 2) = [(3+6) + (3+6+10) + (3+6+10+15)]/70 = 62/70.$$

$$\Pr(\# \geq 3) = [(2+3) + (2+3+4) + (2+3+4+5)]/70 = 42/70.$$

$$\Pr(4) = [(2+3) + (2+3+4)]/70 = 14/70. \text{ Therefore, by Lemma 3 we have}$$

$$A_4 = \frac{1}{4} \sum_{r=1}^4 \Pr(\# \geq r) = [(69/70) + (62/70) + (42/70) + (14/70)]/4 = 187/280.$$

#### A.5 The case of different means

We assume that sellers are randomly drawn from a uniform distribution in the unit interval ( $\text{unif}[0,1]$ ), while the buyers are random draws with a uniform distribution in the interval of  $[\delta, 1 + \delta]$ . We only consider the

case of  $V = U = n$ .

For  $n = 1$ , straightforward calculation shows that  $A_1 = \Pr(1) = (1/2)(1 - \delta)^2$ .

For  $n = 2$ , a more tedious calculation shows that

$$\Pr(0) = (1 + 4\delta + 6\delta^2 - 4\delta^3 - \delta^4)/6, \text{ and } \Pr(2) = (1 - \delta)^4/3.$$

Hence (using  $\Pr(1) = 1 - \Pr(0) - \Pr(2)$ ),

$$A_2 = (1/2)[\Pr(1) + 2\Pr(2)] = (7/12)(1 - \delta)^2 + (1/12)\delta(1 - \delta)^2(2 + 3\delta).$$

## A.6 The case of $V \neq U$

### Case (i) $(U, V) = (1, 2)$ or $(2, 1)$

Let  $v_1 < v_2$  be the order statistic of openings. By Lemma 1 we have

$$\Pr(1) = \Pr(v_2 > u > v_1 \text{ or } u > v_1, v_2) = \Pr(v_2 > u > v_1) + \Pr(u > v_2 > u_1) = 2\{1!2!/3!\} = 2/3.$$

Therefore,  $M_{U,V} = M_{1,2} = \Pr(1) = 2/3$ . From this one can compute  $B_{U,V}$ .

### Case (ii) $(U, V) = (3, 1)$ or $(1, 3)$

Let  $v_1 < v_2 < v_3$  be the order statistic of openings.

$$\Pr(0) = \Pr(u < v_1, v_2, v_3) = \Pr(u < v_1 < v_2 < v_3) = \{1!3!/4!\} = 1/4.$$

$\Pr(1) = 1 - \Pr(0) = 3/4$ . Therefore,  $M_{1,3} = \Pr(1) = 3/4$ . Then one can compute  $B_{U,V}$ .

### Case (iii) $(U, V) = (2, 4)$ or $(4, 2)$

Let  $u_1 < u_2$  and  $v_1 < v_2 < v_3 < v_4$  be the order statistics of candidates and openings, respectively.

$$\Pr(0) = \Pr(u_2 < v_2) = \Pr(u_1 < u_2 < v_1 < v_2 < v_3 < v_4) = \{2!4!/6!\} = 1/15.$$

$$\Pr(1) = \Pr(v_1 < u_1 < u_2 < v_2) + \Pr(v_1 < u_1 < v_2) = 1/15 + 4/15 = 1/3.$$

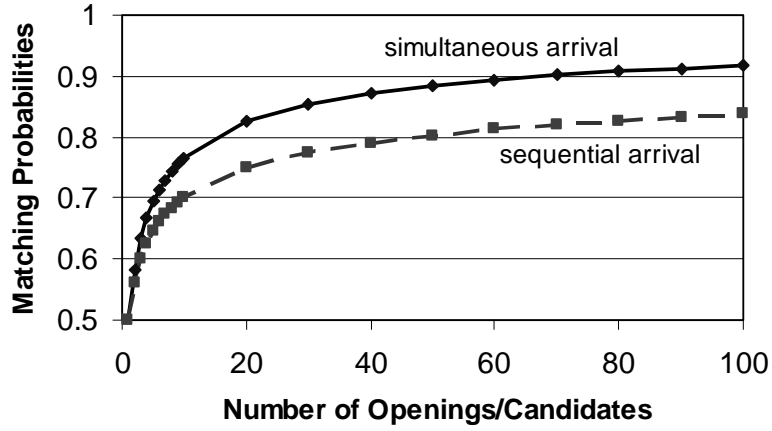
$$\Pr(2) = 1 - \Pr(0) - \Pr(1) = 3/5. \text{ Hence, } M_{2,4} = 1/3 + 2(3/5) = 23/15.$$

## References

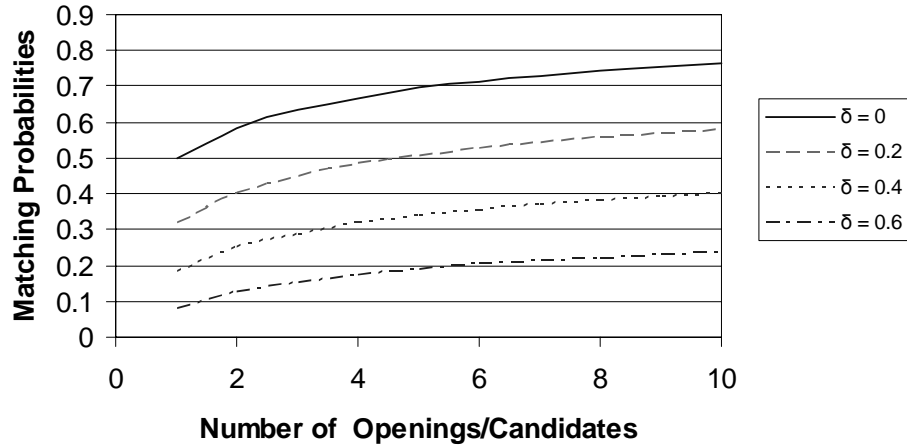
- Berman, E.: 1997, Help wanted, job needed: estimates of a matching function from employment service data, *Journal of Labor Economics* **15**(1), S251–S292.
- Billingsley, P.: 1995, *Probability and Measure*, New York, New York: John Wiley & Sons.
- Blanchard, O. and Diamond, P.: 1989, The Beveridge Curve, *Brookings Papers on Economic Activity* (1), 1–60.

- Blanchard, O. and Diamond, P.: 1994, Ranking, unemployment duration, and wages, *Review of Economic Studies* **61**, 417–34.
- Burdett, K., Shi, S. and Wright, R.: 2001, Pricing and matching with friction, *Journal of Political Economy* **109**(5), 1060–85.
- Chordia, T., Roll, R. and Subrahmanyam, A.: 2000, Commonality in liquidity, *Journal of Financial Economics* **56**, 3–28.
- Coles, M. and Smith, E.: 1998, Marketplace and matching, *International Economic Review* **39**(1), 239–54.
- Diamond, P.: 1982, Aggregate demand management in search equilibrium, *Journal of Political Economy* **90**(5), 881–94.
- Dusansky, R. and Vernon, C.: 1998, Rankings of U.S. economics departments, *Journal of Economic Perspective* **12**(4).
- Howitt, P. and McAfee, P. R.: 1987, Costly search and recruiting, *International Economic Review* **28**(1), 89–107.
- Howitt, P. and McAfee, R. P.: 1988, Stability of equilibria with externalities, *Quarterly Journal of Economics* **103**(2), 261–77.
- Lagos, R.: 2000, An alternative approach to search frictions, *Journal of Political Economy* **108**(5), 851–73.
- Lippman, S. and McCall, J.: 1986, An operational measure of liquidity, *American Economic Review* **76**(1), 43–55.
- Niederle, M. and Roth, A.: 2003, Unraveling reduces mobility in a labor market: gastroenterology with and without a centralized match, *Journal of Political Economy* **forthcoming**.
- Petrongolo, B. and Pissarides, C.: 2001, Looking into the black box: a survey of the matching function, *Journal of Economic Literature* **34**, 390–431.
- Roth, A.: 1984, The evolution of the labor market for medical interns and residents: a case study in game theory, *Journal of Political Economy* **92**(6), 991–1016.

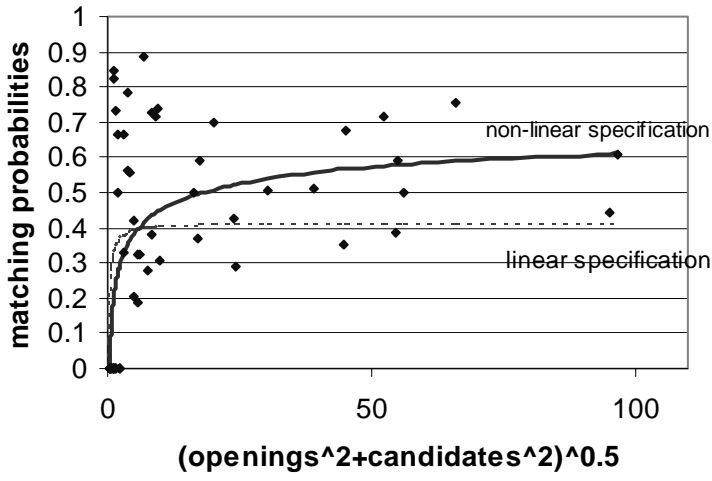
**Figure 1 Matching Probabilities as a Function of Thickness**  
(openings = applicants)



**Figure 2 Matching Probabilities**  
(openings are from uniform(0,1) and candidates are from uniform( $\delta$ , 1))  
(openings = applicants)



**Figure 3a January 2000 Job Market**



**Figure 3b January 2001 Job Market**

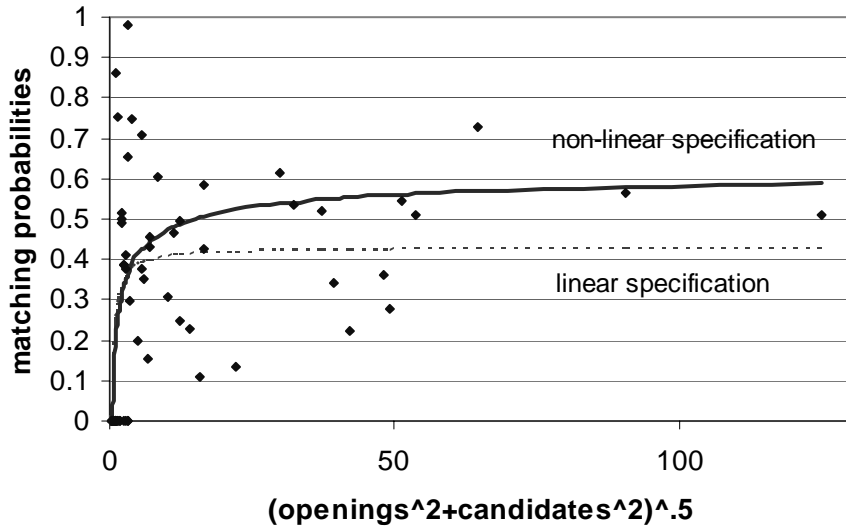


Figure 3c Job Markets 2000 and 2001

