

NBER WORKING PAPER SERIES

EMPIRICAL MATCHING FUNCTIONS:  
ESTIMATION AND INTERPRETATION  
USING DISAGGREGATE DATA

Patricia M. Anderson  
Simon M. Burgess

Working Paper No. 5001

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
January 1995

We would like to thank Hugh Courtney for providing us with regional vacancy data, and John Budd, Al Gustman, Andrew Oswald and participants of numerous seminars for helpful comments. This paper was started while the second author was a visitor at Dartmouth College. This paper is part of NBER's research program in Labor Studies. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

© 1994 by Patricia M. Anderson and Simon M. Burgess. 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.

NBER Working Paper #5001  
January 1995

EMPIRICAL MATCHING FUNCTIONS:  
ESTIMATION AND INTERPRETATION  
USING DISAGGREGATE DATA

ABSTRACT

In this paper, we estimate matching functions using disaggregate data. We find strong support for the matching approach, with most specifications implying slightly increasing returns to scale. This finding does not appear to arise from our inclusion of additional controls or from the level of disaggregation, and so we conclude that earlier findings of constant returns in the US may be due to the various approximations needed to construct an aggregate time series. We also find evidence of endogenous job competition between the employed and nonemployed, so that the estimated parameters from a matching function cannot be interpreted as structural parameters.

Patricia M. Anderson  
Department of Economics  
Dartmouth College  
6106 Rockefeller  
Hanover, NH 03755-3514  
and NBER

Simon M. Burgess  
Department of Economics  
University of Bristol  
8 Woodland Road  
Bristol BS8 1TN  
UNITED KINGDOM

## 1. Introduction

An increasingly popular tool for analyzing the labor market is the search and matching approach, which has now accumulated a substantial literature. While there has been some empirical work using aggregate gross flows data, the majority of this work has been theoretical.<sup>1</sup> This emphasis on theory is well illustrated by the papers in a recent symposium on job matching published in the *Review of Economic Studies* (1994), not one of which present any empirical evidence. This is reinforced by Hall's (1989) comment on Blanchard and Diamond (1989) that what this literature currently lacks is disaggregate evidence. This paper works toward filling that gap by using a unique micro-level dataset with over 15 million quarterly observations on job matches over time, covering all industries. We use these wage records to create quarterly state-industry observations, which we then combine with information on unemployment and vacancies, information on the demographic structure of the labor markets, and relevant features of the unemployment insurance (UI) system in each state. We then use this state-industry panel to estimate matching functions at various levels of aggregation. Additionally, we can separately identify flows from employment to employment and flows from nonemployment to employment. Thus, we also estimate matching functions disaggregated by the source of the hire. We then briefly discuss the importance of employed job searchers in the job matching process, and empirically test for their impact. The results from this exercise highlight the fact that the commonly estimated parameters of a matching function are in fact a reduced form combination of a structural matching function and a job competition model. Thus, care must be taken in interpreting empirical matching functions.

---

<sup>1</sup>See for example, Diamond and Maskin (1979), Hall (1979), Diamond (1982), Mortensen (1982), Pissarides (1985), and Hosios (1990) for some of the theoretical foundations, and Blanchard and Diamond (1989, 1990), Pissarides (1986) and Layard, Nickell and Jackman (1991) for aggregate empirical work.

The remainder of the paper is organized as follows. In section 2 we present the results from estimating matching functions at different levels of aggregation, while section 3 discusses the effect of endogenous job competition. Section 4 then tests the model empirically and discusses the implications of the results for the interpretation of empirical matching functions. Section 5 offers some conclusions.

## 2. Matching Functions

The context for matching functions is a market with incomplete information and trading frictions, exemplified by the absence of an auctioneer (see Diamond and Maskin, 1979, Hall, 1977, Diamond, 1982, Pissarides, 1985). Without this central market coordinating mechanism, agents wishing to trade have to search for a trading partner. The matching function is a reduced form characterization of the number of matches formed in such an environment. It relates the number of searchers on either side of the market (in our case, vacancies and job searchers) to the flow of new matches formed (in our case, new hires). Thus, we have  $M = M(J, V, \mu)$ , where  $M$  is new matches,  $J$  is job searchers,  $V$  is vacancies and  $\mu$  is a set of other influences on the efficiency of the matching process.  $M$  is increasing and concave in both  $J$  and  $V$ , and  $M(0, V, \mu) = M(J, 0, \mu) = 0$ . Typically, a log linear form is assumed in empirical work. One critical question is the degree of homogeneity or returns to scale of the function  $M(\cdot)$ . If  $M(\cdot)$  has increasing returns, there is the possibility of multiple Pareto rankable equilibria in the market. It is likely that other factors, in addition to simply the numbers on either side of the market, will affect the number of matches formed. These are included in  $\mu$ , and might include demographic factors (the young tend to invest more in search than the old), available job search channels (for example, unions may add another method of finding jobs), incentives for search and so on.

Empirical models of the matching process have almost all utilized aggregate time series data.<sup>2</sup> These include Blanchard and Diamond (1989) for the USA, Pissarides (1986) for the UK, and van Ours (1991) for the Netherlands, all of whom use their estimates of the matching function to form the equilibrium relationship between unemployment and vacancies (the Beveridge curve) and hence to analyze the dynamic evolution of unemployment. Each of these studies assumes that only the unemployed are engaged in the matching process (i.e., assumes  $J \equiv U$ , the unemployed) and essentially only includes time trends on the right hand side other than  $U$  and  $V$ . Blanchard and Diamond estimate the elasticities to be around 0.6 for vacancies and 0.4 for unemployment for US aggregate data; Pissarides also finds approximately constant returns to scale for the UK.

Our data allow us to present panel evidence on matching functions. The data are fully described in the Data Appendix, but briefly they derive from the CWBH (Continuous Wage and Benefit History) project in six states over the period 1978 to 1984<sup>3</sup>. This provides quarterly wage records on a sample of the state's covered workers. The worker's record includes a firm identifier and a number of firm characteristics. On the basis of these data, we can discover the number of new matches formed in a state in each industry in each quarter. A new hire is identified when an employer-employee match first appears (other than the first quarter of the sample). We can also separately identify recalls from temporary separations. Merging this with state-level data on unemployment and vacancies (again, see the appendix for details), we estimate a matching function. The vacancy data are a quarterly MSA-level help-wanted advertising series<sup>4</sup>; the unemployment data are taken straight from various issues of *Employment and Earnings*.

---

<sup>2</sup>An exception is Coles and Smith, 1992 and 1994, who investigate cross section matching functions for England.

<sup>3</sup> We thank Bruce Meyer for originally obtaining this dataset.

<sup>4</sup> We are extremely grateful to Hugh Courtney for collecting and providing us with this data. Details can be found in Courtney (1991).

Table 1 presents the results from estimating a matching function using panel data on state-industry level matches, as well as for matches aggregated to the state level. Note that Appendix Table 1 provides definitions of the industry groups used. Paralleling the aggregate literature, we include separate intercepts and trends for each of the state-industry groups (or states for the more aggregated models). We are also able to include a number of other control variables, based on the firm characteristics in the CWBH and other mapped in data. These comprise the change in log employment for the national industry, the log industry share of employment in the state, the average potential weekly UI benefit for the state, and state-level demographic variables (the proportions of white, female, young and old employees) and a measure of unionization. The first two are measures of industry demand (to complement the state vacancy rate), and are only included in the state-industry level equations. The UI benefit is used as a measure of the subsidy to unemployment, while the last set account for differences in search intensity between different groups. Means and full descriptions of all the variables can be found in Appendix Table 2.

The first thing to notice in Table 1 is that in all cases the coefficients on both unemployment and vacancies are significantly positive, as would be implied by a theory of matching. Given that Blanchard and Diamond (1989) note that "one may legitimately wonder whether such a function [a correlation between unemployment, vacancies and new hires in the US aggregate time series data] exists at all." (p. 3), our finding of just such a relationship using disaggregated data seems quite significant. There are differences, however, between our results and past estimates. As seen in model (1), estimates for a disaggregate matching function which includes additional control variables imply somewhat higher returns to scale than are typically found in the simplest aggregate models.<sup>5</sup> In particular, the estimated elasticity with respect to vacancies of 0.83 is about a third larger

---

<sup>5</sup>Instrumented versions of an aggregate matching function, as in Blanchard and Diamond (1989), tend to also imply somewhat higher returns to scale than do the OLS versions.

than has typically been found using aggregate data, while the elasticity with respect to unemployment, at 0.45, is only slightly larger.

Our finding of slightly increasing returns to scale does not, however, seem to be due either to our inclusion of additional control variables, or to our use of disaggregated data. When only state-industry intercepts and trends are included, as in model (2), the returns to scale actually increase slightly. This increase is due to entirely to an increase in the elasticity with respect to unemployment, though, as the elasticity with respect to vacancies is essentially unchanged. To investigate the effect of aggregation, we sum up all matches across industries in the state, creating a state panel. Models (3) and (4) then parallel the specifications of models (1) and (2) using this more aggregated data. When we include state demographic variables the results are similar to the more disaggregated model, with coefficients on the unemployment and vacancy variables of 0.43 and 0.81, respectively. The impact of the additional control variables is even greater at this more aggregated level, though, since in model (4) the elasticities with respect to both unemployment and vacancies increase. Overall, then, our finding of increasing returns to scale is robust to both changes in specification and changes in the level of aggregation.

Thus, it appears that the difference between our results and past work for the U.S. is due to inherent data differences. Given the fact that these differences are large, such a result is not surprising. First, we have data from only four states, which may not be representative of the U.S. as a whole.<sup>6</sup> Perhaps more important, though, there are major differences in how job flows are calculated in our data versus that of Blanchard and Diamond (1989). Since our data is at the level of the job match, we are able to quite accurately identify the number of new hires. By contrast, Blanchard and Diamond must rely on a number of ad hoc adjustments to obtain this number. They begin by using the Current Population Survey (CPS) to calculate flows into employment from unemployment

---

<sup>6</sup>Anderson and Meyer (1994) compare these states, plus four others, to the U.S., finding that aggregate growth was slightly slower.

and out of the labor force. Following Abowd and Zellner (1985), this number is adjusted to reflect misreporting.

More significant adjustments involve the addition of employment to employment flows and the removal of recalls. In order to obtain the first of these, they assume that the aggregate quit rate is identical to that in manufacturing, and that employment to employment quits are fixed at 40 percent of total quits. Given that manufacturing turnover is quite a bit below total turnover (see Anderson and Meyer (1994) for example), such assumptions may be suspect. Aggregate recalls are assumed to be 1.5 times manufacturing recalls, although no hard evidence is presented to indicate that two-thirds of all recalls occur in manufacturing. Finally, their data is seasonally adjusted, while ours is not.

Evidence presented in Table 1 of Blanchard and Diamond (1989) indicates that their results are particularly sensitive to the definition of employment to employment flows. The returns to scale drop from 0.90 to 0.76 when flows are defined as twenty percent of quits, with this change due mainly to a fall in the elasticity with respect to vacancies. Perhaps even more important than the level chosen for the fraction of quits which are deemed to be job-to-job flows is the fact that this fraction is assumed to be fixed. It may well be that the true fraction fluctuates with the business cycle. If this is the case, then by using a fixed fraction, Blanchard and Diamond are leaving out a component of the dependent variable which varies positively with vacancies, and their measure will result in lower coefficients than would the true measure. Overall, then, consideration of all these differences implies that their smaller elasticity may be due to their mismeasurement of employment to employment flows.

We also present the results for the other right hand side variables in Table 1. The employment growth rate of the industry at the national level has a positive effect as expected and is strongly significant. The industry employment share in the state is also positive, but not significant. The average dollar value of potential weekly UI benefits for a state-industry group is significantly negative, indicating that higher UI benefits slow down



the rate of matching for given levels of unemployment and vacancies. Again, this is as expected since an increase in the value of unemployment will increase the reservation wage. The table also shows that a high proportion of female, young (under 25 years), old (over 55 years), nonwhite or union workers raises matching efficiency, but only the first two are even marginally significant. These results are in accord with the idea that the most stable workers are prime age white males. Thus, we see that high proportions of workers from groups which are more likely to be highly mobile, such as the young, are associated with more matches.

An additional advantage of data which are disaggregated at the industry level is the ability to estimate separate matching functions for different industries. In Table 2 we estimate analogs of model (3) from Table 1 (which is repeated here as model (1) for convenience), where we have summed matches in a state for major industry groups. While the standard errors are fairly large, some major differences across industry group are apparent. Most notable, perhaps, is the high level of returns in Construction. Here we see elasticities with respect to unemployment and vacancies of 1.62 and 2.65 respectively. At the other extreme, the comparable figures for Retail Trade are just 0.29 and 0.55. All other industries exhibit at least slightly increasing returns to scale, except Manufacturing. For Manufacturing, returns are almost exactly constant.

While it is difficult to pinpoint the source of these cross-industry differences, we can speculate on some possibilities. The large elasticities found in Construction seem likely to be a reflection of the unique character of this industry, in which each new building project implies a new workforce. Also clearly important will be the different proclivities of different industries to advertise job openings in newspapers (recall that our vacancy measure is based on total help-wanted ads placed in local newspapers). For example, the lower elasticity in retail trade may reflect a greater tendency for shop owners to advertise openings through the use of window signs. Thus, one needs to keep in mind that the differences found across industries are due not only to true differences in matching, but

also to data limitations. Differences in the elasticities with respect to unemployment are also apparent, but perhaps less easily explained. To the extent, though, that an industry exhibits countercyclical tendencies, it is likely to be less responsive to state unemployment.

Overall, then, the matching function appears to emerge from this estimation as a clear empirical fact: there is a strong correlation in the microdata between hires, unemployment and vacancies. However, before we pronounce the matching approach to be supported by the data, we highlight a potentially serious problem, which our data will also allow us to address. In our data we can identify the origin of workers forming new matches, that is, we can determine whether they have come directly from another job or not. Thus, in Table 3, we repeat the main specification from Table 1, along with the corresponding matching functions<sup>7</sup> for hires of the nonemployed<sup>8</sup> and of the employed<sup>9</sup>. What we find is that the key coefficients on unemployment and vacancies are quite different between these groups. Both coefficients are considerably larger for job-to-job matches, although the proportional change in the elasticity with respect to unemployment is greater. There are also differences in the effect of the demographic variables. Most noticeably, none of the state demographic shares are significant in model (2), while all of them are significant in model (3). This finding seems to reinforce our interpretation that these variables are capturing differences in job mobility among these groups.

In the remainder of this paper, we offer an explanation of the results in Table 3, arguing that gross flows of workers into jobs are governed by two processes: job matching and endogenous job competition. The latter describes the idea that not only are the unemployed looking for jobs, but that a significant (and endogenous) fraction of the employed are engaged in job search. Thus, the fraction of the new matches won by these

---

<sup>7</sup> Whether these functions should be called matching functions is a moot point at this stage.

<sup>8</sup> We cannot separate out the unemployed and out-of-the-labor-force states in our data.

<sup>9</sup> Obviously, only two of these three columns are independent; we present all three for convenience.

two groups depends on the state of the labor market. As a result, the reduced form relationships between hires of the unemployed and hires of the employed with unemployment and vacancies differ significantly from the overall matching function.

### 3. A Theoretical Model of Endogenous Job Competition

The main feature that we add to the standard job matching model is the idea of endogenous job competition (see Burgess, 1993, for more details), arguing that this is an essential component to an understanding of the gross worker flows in the labor market. The central point is very simple. Employed workers monitor the state of the labor market, and when the economy is booming and there are a lot of job offers available, more of the employed find it profitable to spend time searching for a better job. However, their entry into the labor market reduces the probability of any one searcher, employed or unemployed, finding a job. This effect dampens the impact of changes in hiring or vacancies on the flow out of unemployment. Below we outline a simple theoretical model and then formulate an empirical version that we can apply to our data.<sup>10</sup>

#### A BASIC MODEL

While definitely a minority interest, search theory has long recognized the fact that the employed engage in job search (see Burdett, 1978 Mortensen, 1986) Under the standard conditions, each of the employed calculate a reservation wage, and if she earns less than this wage she engages in job search. This reservation wage depends on the arrival rate of job offers,  $\theta$ , and on all other parameters of the search problem, including

---

<sup>10</sup>There has of course been a lot of work on employee quits, with recent examples being McLaughlin (1991) and Farber (1994). Additionally, evidence on worker separations based on the data used in this paper is available in Anderson and Meyer (1994). However, this section is concerned neither with total quits nor total separations, but rather with job-to-job transitions as viewed from the receiving firm (rather than from the firm being left).

personal characteristics of the worker, denoted  $Z_i$ . The proportion of the employed who are searching for new jobs is then simply the proportion earning less than this wage, denoted  $\phi(\theta, Z)$ ,  $0 \leq \phi \leq 1$ , here  $Z$  represents the distribution of  $Z_i$  through the working population. It is the endogeneity of  $\phi$ , through its dependence on  $\theta$ , which is critical to the analysis which follows.

The first question we therefore have to consider is the determination of  $\theta$ . Since the decision of the employed job searchers is based on *their* arrival rate of offers, we will need to consider the link between that and the offer rate for the unemployed job searchers. This will depend chiefly on the preferences of employers, that is, on the degree of substitutability they perceive between the two groups. One way of expressing this is in terms of ranking, which describes the degree to which one group is preferred to the other.<sup>11</sup> Unfortunately, we will not be able to directly estimate the degree of ranking, because  $\phi$  is unobservable. To see why this is so, suppose there are 200 people currently unemployed and suppose that of 100 new hires, 50 come from the ranks of the unemployed and 50 from the employed. Because we do not know how many of the employed are engaging in search (we do not know  $\phi$ ), we cannot say whether this arises from 200 employed job searchers and "equal chances" (no ranking), or 100 employed job searchers and an employer preference for employed applicants. What we do here is take two different assumptions and work out the implications for an empirical model of the relationship between total new matches and the matches of the nonemployed. Note that our ability to test the endogenous job competition hypothesis, namely the relationship of  $\phi$  to  $\theta$  remains unaffected.

First, we assume that employers have no preference for one group over the other and hence the probability of receiving a job offer is the same for an unemployed searcher as an employed searcher. The probability of receiving a job offer is then given by  $\theta = M/J$

---

<sup>11</sup>Blanchard and Diamond, 1994, use this to describe employer preferences between short-term and long-term unemployed job applicants

where  $M$  is the number of matches made, and  $J$  is the number of job searchers,  $J = U + (L - U)\phi$ ,  $U$  being the nonemployed job searchers, and  $L$  the labor force. We now have two equations in  $\theta$  and  $J$ , jointly determined by  $U$ ,  $M$  and  $Z$ . Solving these gives:

$$(1) \quad \theta = \theta(M, U, Z), \text{ and } J = J(M, U, Z).$$

The two key worker flows processes are the matching function and the competition between unemployed and employed workers for the new matches. We could parameterize the latter by estimating either an unemployed transition equation or an employed transition equation or an equation explaining the shares of accessions won by the nonemployed. We choose this final option on the grounds that it seems a natural format, provides a convenient normalization and yields a straightforward test of the endogenous job competition hypothesis. Define  $\beta = X/M$ , where  $X$  is the actual number of transitions into employment from nonemployment, and  $\beta$  can be expressed as:

$$(2) \quad \beta = \theta(M, U, Z) * (U/M) * \rho,$$

where  $\rho$  is the relative offer acceptance rates of the two groups.

The elasticities of  $\beta$  with respect to  $M$  and  $U$  are:

$$(3) \quad \varepsilon_{\beta, M} = \varepsilon_{\theta, M} - 1 = \frac{-(1 - \pi)\eta}{1 + (1 - \pi)\eta} < 0$$

$$(4) \quad \varepsilon_{\beta, U} = \varepsilon_{\theta, U} + 1 = \frac{1 + (1 - \pi)\eta - (1 - \phi)\pi}{1 + (1 - \pi)\eta} > 0$$

Where  $\eta$  is the elasticity of  $\phi$  with respect to  $\theta$  and  $\pi$  is the fraction of searchers who are nonemployed, and the elasticity with respect to  $U$  is conditional on  $M$ , since this is what we will need in the empirical work. Note that if  $\eta=0$ , then  $\varepsilon_{\beta, M} = 0$ .

If we write a general matching function of the form  $M(J, V, \mu)$  and utilize  $J(M, U, Z)$  to substitute out  $J$ , then we reach the standard matching equation relating total hires to unemployment and vacancies, estimated by Blanchard and Diamond (1989) and others. It is now clear that such an equation is not a structural matching function, but rather a reduced form combination of the matching function proper and the endogenous job

competition function (1). In particular, the parameters estimated from such a function cannot be identified directly with those of the true matching function.

#### A RANKING MODEL

We now consider an alternative assumption about the nature of competition between employed and unemployed job searchers, taking an equally idealized case where employers have a marginal preference for employed applicants.<sup>12</sup> Such an assumption is not difficult to rationalize, given that the very fact of being employed shows that at least one employer finds the applicant profitable to hire. Thus, there may be a view that hiring from the ranks of the nonemployed poses a significantly greater risk of picking a lemon. This is similar in spirit to the results of Gibbons and Katz (1991) who find that workers displaced by a plant closing are less stigmatized than those workers who are laid off for other reasons. Also supportive of this assumption is the finding by Blau and Robins (1990) that the employed generate more offers per search than the unemployed. Here we characterize employer's preference as being such that if a vacancy is visited by some employed and some unemployed applicants, then the job will always be offered to one of the former. However, the preference is not strong enough to justify a firm turning away an unemployed applicant in the hope of finding an employed worker next period. This means that a vacancy will be in one of four states: it may have no applicants, it may have just unemployed applicants, or just employed applicants, or some applicants of each type. In the last two cases, an employed worker will be offered the job. There is now an asymmetry in the interaction of the two types of agents. Employed job searchers are entirely unaffected by changes in the number of unemployed job searchers, conditional on the number of job offers,  $M$ . Of course  $\phi$ , and hence  $\beta$ , will still be influenced by  $M$ . So

---

<sup>12</sup>This follows an idea in Blanchard and Diamond (1994), who assume that employers rank short duration unemployed applicants ahead of the long-term unemployed.

the effect of  $U$  on  $\beta$  comes entirely through its influence on  $M$ . The success of unemployed job searchers is still affected by the number of employed searchers.

This can be briefly formalized as follows: Let  $\theta^N$  be the job offer rate for the employed, so that

$$(5) \quad \theta^N = \bar{\theta}(M, \phi N)$$

$$(6) \quad \phi = \bar{\phi}(\theta^N, Z)$$

Equation (8) captures the idea that the job chances of the employed depend only on their numbers and the number of offers (not, as before, on the total number of job searchers).

These two equations combine to give  $\theta^N(M, N, Z)$ . Noting that the number of matches for the employed is  $\theta^N N \phi$ , and recalling the definition of  $\beta$ , we have:

$$(7) \quad \beta = 1 - \frac{\theta^N N \phi}{M}.$$

Hence  $\beta$  is independent of  $U$ , conditional on  $M$ ,<sup>13</sup> and the elasticity of  $\beta$  with respect to  $M$

is:

$$(8) \quad \varepsilon_{\beta, M} = -\left(\frac{1-\beta}{\beta}\right)\eta < 0.$$

The main point to take from this section is not so much this specific result from the particular assumption used, but rather the following general idea. The less effective competition the unemployed represent for the employed job searchers, the more insulated will the latter be from changes in the numbers unemployed. Thus, while we cannot directly estimate the degree of ranking we would expect  $\varepsilon_{\beta, U}$  to increase from zero to the expression in (4) as the degree of ranking falls.

#### 4. Estimating an Empirical Model of Job Competition

---

<sup>13</sup>It is independent in a behavioral sense. If, as above, we replace  $N$  by  $L-U$ , then there is a minor, purely mechanical, dependence.

In order to empirically implement the theory, we simultaneously estimate a model of job competition along with the matching function of column (1) from Table 1, using two stage least squares (2SLS). Four lags of  $M$  are included in the instrument list. The job competition equation includes  $M$  and  $U$ , which are the key coefficients we focus upon. We also include the set of state demographic variables discussed above, as well as separate intercepts and trends for each state-industry group, along with two additional variables measuring the state-industry average firm size and per capita wage bill. The latter two may pick up size and wage effects in hiring, as well as determinants of  $p$ .

#### RESULTS FOR THE JOB COMPETITION MODEL

Table 4 presents the results from estimating the job competition model specified above. As predicted by the theory, the estimated coefficient on accessions is significantly negative. This is the key test of the endogenous job competition hypothesis. The coefficient on unemployment is essentially zero. This fits the ranking model better than the no ranking model. The effects of the other included variables follow a pattern similar to that found for the matching functions. Here, the proportions female and young each have a significantly negative effect on  $\beta$ , and now the effect of union is also significant. It is reasonable that we find coefficients of the opposite sign to those found in Table 1, since with increased matching efficiency, workers are more likely to move quickly from job to job, with no observable spell of unemployment.

#### IMPLICATIONS FOR MATCHING MODELS

The evidence found above for the role of endogenous job competition has implications for matching models. For example, the relationship among the elasticities estimated across the three columns of Table 3 can now be interpreted in light of the endogenous job competition model. Letting  $M$  refer to total new hires,  $X$  to new hires from nonemployment, and  $Q$  to new hires from employment, and given that  $\ln \beta = \ln X -$



In  $M$ , we have the following expressions for elasticities with respect to vacancies and unemployment:

$$(9) \quad \varepsilon_{X,V} = \varepsilon_{M,V} \varepsilon_{X,M} = (1 + \varepsilon_{\beta,M}) \varepsilon_{M,V}$$

$$(10) \quad \varepsilon_{Q,V} = \varepsilon_{M,V} \left( 1 - \left( \frac{\beta}{1-\beta} \right) \varepsilon_{\beta,M} \right)$$

$$(11) \quad \varepsilon_{X,U} = \varepsilon_{M,U} \varepsilon_{X,M} + \varepsilon_{\beta,U} = \varepsilon_{M,U} (1 + \varepsilon_{\beta,M}) + \varepsilon_{\beta,U}$$

$$(12) \quad \varepsilon_{Q,U} = \varepsilon_{M,U} \left( 1 - \left( \frac{\beta}{1-\beta} \right) \varepsilon_{\beta,M} \right) - \left( \frac{\beta}{1-\beta} \right) \varepsilon_{\beta,U}$$

The negativity of  $\varepsilon_{\beta,M}$  and the finding that  $\varepsilon_{\beta,U}$  is approximately zero are sufficient to generate the results in Table 3. In words, the sensitivity of the employed job searchers to the offer probability means that the returns to scale in a function relating job-to-job flows to  $U$  and  $V$  will be greater than in a function relating nonemployment-to-jobs flows to  $U$  and  $V$ .

## 5. Conclusions

In this paper we provide an empirical study of a central feature of much recent labor and macroeconomic research: the matching function. Using high quality data based on individual worker-firm match records, we establish a correlation (across states, industries and time) between the number of new matches formed, unemployment and vacancies. The estimates imply increasing returns to scale in the matching function. This does not appear to arise from our inclusion of additional controls or the from the level of disaggregation, and we conclude that earlier findings of constant returns in the US may be due to the various approximations needed to construct an aggregate time series.

The matching function appears to be an important component of an understanding of gross worker flows into jobs. Another component is the competition between employed and unemployed job searchers for the new positions. The central idea here is the endogeneity of the fraction of the employed who participate in the search for new jobs.

Our evidence suggests that this idea has support in the data. We also tentatively find support for the view that employers prefer employed to unemployed applicants. If the endogenous job competition is true, then this implies that correlations between hires, unemployment and vacancies must be interpreted as reduced form relationships, not structural matching functions.

## References

- Anderson, Patricia M. and Bruce D. Meyer. "The Extent and Consequences of Job Turnover." *Brookings Papers on Economic Activity, Microeconomics* (1994): 177 - 248.
- Blanchard, Olivier J. and Peter A. Diamond. "The Beveridge Curve." *Brookings Papers on Economic Activity* 1 (1989): 1 - 60.
- \_\_\_\_\_. "The Cyclical Behavior of the Gross Flows of U.S. Workers." *Brookings Papers on Economic Activity* 2 (1990): 85 - 156.
- Blank, Rebecca M. and David E. Card. "Recent Trends in Insured and Uninsured Unemployment: Is There an Explanation?" *Quarterly Journal of Economics* 56 (November 1991): 1157 - 1190.
- Blau, David M. and Philip K. Robins. "Job Search Outcomes for the Employed and Unemployed." *Journal of Political Economy* 98 (June 1990): 637 - 655.
- Burdett, Kenneth. "Employee Search and Quits." *American Economic Review* 68 (March 1978): 212 - 220.
- Burgess, Simon M. "A Model of Competition between Employed and Unemployed Job Searchers: An Application to the Unemployment Outflow Rate in Britain." *Economic Journal* 103, (September 1993): 1190 - 1204.
- Courtney, Hugh. "The Beveridge Curve and Okun's Law: A Re-examination of Fundamental Macroeconomic Relationships in the US." unpublished Ph. D. dissertation. Cambridge, MA: M.I.T., 1991.
- Diamond, Peter A. "Aggregate Demand Management in Search Equilibrium." *Journal of Political Economy* 90 (October 1982): 881 - 894.
- Diamond, Peter A and Eric Maskin. "An Equilibrium Analysis of Search and Breach of Contract, I: Steady States." *Bell Journal of Economics* 10 (Spring 1979): 282 - 316.
- Ehrenberg, Ronald G. and Robert S. Smith. *Modern Labor Economics*. New York: Harper Collins, 1991.
- Gibbons, Robert and Lawrence Katz. "Layoffs and Lemons." *Journal of Labor Economics* 9 (October 1991): 351 - 380.
- Gottschalk, Peter and Tim Maloney. "Involuntary Terminations, Unemployment, and Job Matching: A Test of Job Search Theory." *Journal of Labor Economics* 3 (April 1985): 109 - 123.
- Hall, Robert E. "A Theory of the Natural Rate of Unemployment and the Duration of Unemployment." *Journal of Monetary Economics* 5 (April 1979): 153 - 169.
- Hall, Robert E. "Comment on: Blanchard, Olivier J. and Peter A. Diamond." *Brookings Papers on Economic Activity* 1 (1989): 61 - 64.

- Hosios, Arthur J. "On the Efficiency of Matching and Related Models of Search and Unemployment." *Review of Economic Studies* 57 (April 1990): 279 - 298.
- Layard, Richard, Stephen Nickell and Richard Jackman. "Job Search: the Duration of Unemployment." *Unemployment*. Oxford: Oxford University Press, 1991.
- Mortensen, Dale T. "The Matching Process as a Non-Cooperative Bargaining Game." *The Economics of Information and Uncertainty*, edited by J. J. McCall. Chicago: NBER University of Chicago Press, 1982.
- Mortensen, Dale T. "Job Search and Labor Market Analysis." *Handbook of Labor Economics*, edited by Orley Ashenfelter and Richard Layard. Amsterdam: North Holland Press, 1986.
- Pissarides, Christopher A. "Short-run Equilibrium Dynamics of Unemployment, Vacancies and Real Wages." *American Economic Review* 75 (September 1985): 676 - 690.
- Pissarides, Christopher A. "Unemployment and Vacancies in Britain." *Economic Policy* 3: 499 - 559.
- Troy, Leo and Neil Sheflin. *U.S. Union Sourcebook*. West Orange, NJ: Industrial Relations Data and Information Services, 1985.
- U.S. Department of Labor. *Employment and Earnings*. Washington, DC: U.S. Government Printing Office, various issues 1978 - 1985.
- U.S. Bureau of the Census. *Census and You*. Washington, DC: U.S. Government Printing Office, 1993.
- van Ours, J. C. "The Efficiency of the Dutch Labour Market in Matching Unemployment and Vacancies." *De Economist* 139, (1991): 358 - 378.
- Vroman, Wayne. "The Decline in Unemployment Insurance Claims Activity in the 1980's." *Unemployment Insurance Occasional Paper 91-2*. Washington, DC: U.S. Department of Labor - ETA, 1991.

Table 1  
Estimated Matching Functions

Dependent Variable	state-industry level matches	state-industry level matches	state-level matches	state-level matches
Control Variable	(1)	(2)	(3)	(4)
log of state unemployment	0.445 (0.090)	0.567 (0.087)	0.429 (0.176)	0.542 (0.169)
log of local help-wanted rate	0.825 (0.097)	0.808 (0.079)	0.813 (0.203)	0.973 (0.151)
proportion female	4.152 (1.970)	--	1.100 (3.775)	--
proportion under 25	2.133 (1.181)	--	2.994 (2.255)	--
proportion over 55	3.349 (2.076)	--	-1.196 (4.016)	--
proportion white	-0.873 (0.757)	--	-0.117 (1.440)	--
proportion unionized	0.681 (0.499)	--	1.069 (0.990)	--
potential weekly UI benefit	-0.006 (0.002)	--	-0.044 (0.014)	--
change in log industry employment	1.353 (0.151)	--	--	--
log share of state employment	0.002 (0.028)	--	--	--
Mean of Dependent Variable	3.867	3.867	7.001	7.001
Number of Observations	1060	1060	56	56
Adjusted R <sup>2</sup>	0.902	0.891	0.863	0.843

Notes: Standard Errors in parentheses. Models (1) and (2) also include separate intercepts and trends for each state-industry. Models (3) and (4) also include separate intercepts and trends for each state. All dependent variables are in logs.

Table 2  
Estimated Matching Functions  
for Selected Industries

	Aggregate	Construction	Manu- facturing	Wholesale Trade
log of state unemployment	0.429 (0.176)	1.620 (0.646)	0.343 (0.262)	0.057 (0.398)
log of local help-wanted rate	0.813 (0.203)	2.648 (0.717)	0.665 (0.341)	1.239 (0.421)
Adjusted R <sup>2</sup>	0.863	0.705	0.883	0.602
		Retail Trade	FIRE	Services
log of state unemployment		0.289 (0.240)	1.082 (0.291)	0.693 (0.190)
log of local help-wanted rate		0.549 (0.278)	1.044 (0.315)	1.025 (0.204)
Adjusted R <sup>2</sup>		0.823	0.583	0.732

Notes: Standard Errors in parentheses. All models have 56 observations and include all of the variables of model (3) in Table 1. All dependent variables are in logs.

Table 3  
 Estimated Matching Functions  
 Disaggregated by Source of Hire

Dependent Variable	All New Hires	New Hires from Nonemployment	New Hires from Employment
Control Variable	(1)	(2)	(3)
log of state unemployment	0.445 (0.090)	0.349 (0.097)	0.676 (0.177)
log of local help-wanted rate	0.825 (0.097)	0.675 (0.105)	0.980 (0.191)
proportion female	4.152 (1.970)	-0.703 (2.133)	20.661 (3.911)
proportion under 25	2.133 (1.181)	-1.401 (1.279)	15.874 (2.347)
proportion over 55	3.349 (2.076)	2.328 (2.248)	10.558 (4.090)
proportion white	-0.873 (0.757)	0.438 (0.819)	-4.912 (1.515)
proportion unionized	0.681 (0.499)	-0.534 (0.540)	6.398 (1.025)
potential weekly UI benefit	-0.006 (0.002)	-0.005 (0.002)	-0.010 (0.004)
change in log industry employment	1.353 (0.151)	1.334 (0.163)	1.494 (0.296)
log share of state employment	0.002 (0.028)	0.003 (0.030)	-0.010 (0.057)
Mean of Dependent Variable	3.867	3.501	2.621
Number of Observations	1060	1060	1060
Adjusted R <sup>2</sup>	0.902	0.888	0.714

Notes: Standard Errors in parentheses. All models also include separate intercepts and trends for each state-industry. All dependent variables are in logs.

Table 4  
Estimated Model of Job Competition

Variable	Percent of New Hires from Nonemployment
log of new hires in the state-industry	-0.128 (0.036)
log of state unemployment	0.001 (0.041)
proportion female	-4.022 (1.056)
proportion unionized	-1.089 (0.287)
proportion under 25	-3.033 (0.626)
proportion over 55	-0.449 (1.213)
proportion white	1.185 (0.452)
log of average monthly employment	0.007 (0.010)
log of average per capita wage bill	-0.002 (0.013)
Mean of Dependent Variable	4.240
Number of Observations	1060
Adjusted R <sup>2</sup>	0.210

Notes: Standard Errors in parentheses. Models also includes separate intercepts and trends for each state-industry. Dependent variable is in logs. Model is estimated using two stage least squares, with 4 lags of new hires included in instrument list.



## Data Appendix

We construct a panel of state-industry level accessions, state unemployment and vacancy rates, state level demographic shares, and various other state and/or industry level characteristics by merging together data from a variety of sources. These sources include unemployment insurance administrative records, the Current Population Survey (CPS), and assorted published and unpublished data, all of which are described in more detail below.

### DATA ON TRANSITIONS

Our main data on labor market transitions come from administrative records from the unemployment insurance (UI) systems of 4 states<sup>14</sup> (Georgia, Louisiana, Missouri and South Carolina) which participated in the Continuous Wage and Benefit History (CWBH) project.<sup>15</sup> The data are of two types - quarterly wage records and weekly UI records. The quarterly wage records are for a sample (typically 10 to 20 percent) of the state's covered workers. The main category of noncovered workers is the self-employed, so we have a sample of over 90 percent of those employed in the state. In addition to the dollar amount of wages received by the employee, the record contains a firm identifier, and several firm characteristics, including 4 digit SIC industry, average monthly employment over the quarter, and the total quarterly wage bill. The number of quarters of data available differs by state, but averages about 15 quarters between 1979 and 1984. Since there are over 22 million quarterly wage records available, we draw a random sample of individuals, giving us approximately 100,000 to 200,000 wage records from each state. For this sample of the individual workers we have created quarterly job-match histories by sorting the wage

---

<sup>14</sup>Data on Idaho and New Mexico are also available, but dropped due to lack of appropriate vacancy data. Results are similar if they are included using proxies for vacancies, but with the expected attenuation bias due to measurement error.

<sup>15</sup>We thank Bruce Meyer for originally obtaining this data set.

records by employer identification code and employee. A new hire is then defined if an employer-employee job match first appears in a quarter other than the first quarter of data collection. Analogously, a permanent separation is defined if a job match last appears in a quarter other than the last quarter of data collection. If there is a gap in this quarterly job-match history, the quarter before the gap defines a temporary separation, and the quarter after defines a return from a temporary separation.

This method will, however, miss any temporary layoffs which do not encompass an entire calendar quarter. However, almost half of all job losers who are unemployed for just 3 - 4 weeks apply for UI (Vroman, 1991), and we will be able to identify compensated unemployment after matching the UI experience to the wage records. It is likely, though, that this will capture even more than half of such short temporary layoffs, since job losers who actually return to the previous job are also more likely to have other characteristics associated with higher levels of UI receipt, such as being older and union members (Blank and Card, 1991). Additionally, the empirical work generally excludes recalls from temporary layoff, and thus this problem will not effect these estimates. In order to match the UI experience with the job-match histories, we summarize the weekly UI data into quarterly records and match these by person and quarter of initiation to the wage record sample. If a UI claim is found to be initiated in a quarter not previously coded as a separation, that quarter is assumed to contain a temporary layoff. Since we cannot actually determine if this layoff ended in that quarter or the next, we arbitrarily code the return from temporary layoff for the same quarter.

Having identified that an accession occurred, the next step is to classify whether the worker's previous labor market state was employment or nonemployment. If the worker was not in some job match in the previous quarter, or if the worker received UI, the accession is classified as coming from nonemployment. For workers who were in a job match previously and did not receive UI, we must estimate whether the move was directly from employment or not. To do this, we compare earnings in the quarter prior to the

separation with earnings in the quarter of the accession. The number of "earnings weeks" lost is then defined as the difference between old and new earnings, divided by old earnings over 13. If this estimate of the number of weeks unemployed is greater than 2, we classify the worker as coming from nonemployment. Obviously this process will not be perfect. Most notably, accessions into much better paying jobs which occur after relatively short spells of uncompensated unemployment will be misclassified as accessions from employment, and accessions into poorer paying jobs which occur with no intervening unemployment will be misclassified in the reverse manner. Also, moves from employment in another state with no intervening unemployment will be misidentified as new entrants.<sup>16</sup> However, as discussed below, we can roughly evaluate how our classification compares to other findings in the literature.

As shown in appendix table 1, we define 20 industry groups by splitting some of the larger 1 digit SIC groups into subcategories. Then to obtain the final data set containing quarterly observations for each state-industry cell, we sum total accessions, new hires, total accessions from nonemployment and new hires from nonemployment by state, industry group and calendar quarter. We then define  $\beta$  as the number of new hires from nonemployment over the total number of new hires. Looking at the mean of  $\beta$  in Table 3 of the text shows that we have classified 70 percent of new accessions as coming from nonemployment.<sup>17</sup> While an exactly comparable statistic is not readily available, one can be constructed and then compared to our results. So for example, using the PSID, Gottschalk and Maloney (1985) find that 49 percent of job movers experienced some unemployment. In order to translate this into a number comparable to  $\beta$ , we need to realize that accessions come not only from job movers, but also from labor force entrants. Ehrenberg and Smith (1991) present data from the CPS showing that flows from out of the

---

<sup>16</sup>Interstate moves are actually relatively rare. According to the Census Bureau (1993), between 1990 and 1991, only 2.9% of Americans changed state.

<sup>17</sup>The comparable figure for total accessions is 78 percent.

labor force into employment are 1.7 times those from unemployment into employment. Thus, for every 100 separations followed by reemployment, 49 will lead to accessions from unemployment, based on the results of Gottschalk and Maloney. However, based on Ehrenberg and Smith, an additional 83 accessions will occur from out of the labor force. Thus, of a total of 183 accessions, 132, or 72 percent, will be from nonemployment. Based on this calculation, then, our  $\beta$  appears to compare very well with what is known about such transitions.

#### ADDITIONAL SOURCES OF DATA

To estimate the matching functions, we would ideally like information on quarterly vacancies and unemployment at both the state and state-industry level. Unfortunately, in not all cases is such information readily available. Monthly state unemployment rates are available in various issues of *Employment and Earnings*, so we append the state unemployment rate for the last month of the previous quarter to all industry group observations in a state. Additionally, Courtney (1991) has assembled a set of quarterly, MSA-level help-wanted advertising series.<sup>18</sup> For both Georgia and Louisiana, one city within the state is available. Thus we use the series for Atlanta and New Orleans for all of the observations from Georgia and Louisiana respectively. For South Carolina, we use Charlotte, NC, since parts of SC are included in that SMSA. Data from Kansas City and St. Louis are averaged and used for all observations from Missouri.

Given that these city vacancy rates are proxying for state rates, we compare the labor market conditions of the state and the chosen local area to evaluate how damaging this approximation is likely to be. Appendix Figure 1 plots the unemployment rate for the chosen city against that of the state, using monthly data for two years (November 1990 through November 1991), along with a fitted regression line. The slope of this line is

---

<sup>18</sup>We are extremely grateful to Hugh Courtney for providing us with this data. A more complete description of the series can be found in Courtney (1991).

1.05, with a standard error of 0.06. Thus, while the substitution of MSA level vacancies for state level vacancies is not ideal, in general it is likely that the two are quite highly correlated.

We also use several variables that may affect the search and matching process: the change in log employment for the national industry, the log industry share of employment in the state, the average potential weekly UI benefit for the state-industry, and state demographic shares. The first two are included to reflect industry demand, since as described above, our vacancy measure does not vary by industry. Additionally, state demographics proxy for the fact that search effectiveness is likely to differ across different types of individuals, while UI benefits provide a subsidy to unemployment. We also include the available firm demographics in the  $\beta$  equation. These additional variables for the empirical model are obtained in several ways.

First, the change in log employment is calculated based on quarterly industry employment levels presented in *Employment and Earnings*. The potential weekly UI benefit is derived from the CWBHC wage records by applying the appropriate state benefit schedules to the calculated base period earnings. The CWBHC wage records are also used to compute firm demographics (the average per capita wage bill and average monthly employment) by state-industry group. Demographic shares (proportions female, under 25, over 55, and white) are calculated for each state directly off of the March CPS. Each quarter of the year and each industry within the state are then assigned the same value of these variables.<sup>19</sup> Additionally, we count up the number of observations in a state-industry cell of the CPS and adjust this number by the state-industry unemployment rate to obtain employment counts which can be used to construct industry shares for each state.

Data on unionization rates are obtained from a variety of sources. The *U.S. Union Sourcebook* (Troy and Sheflin 1985) provides state unionization rates for 1980 and 1982.

---

<sup>19</sup>We experimented with using state-industry demographic shares, but the cell sizes were deemed too small for the estimates to be reliable.

Appendix Table 1

Industry Group Definitions

Industry Group Name	SIC Codes
Agriculture	1 - 9
Mining	10 - 14
Construction - General	15
Construction - Heavy	16
Construction - Special Trades	17
Manufacturing - Durables	24, 25, 32 - 39
Manufacturing - Nondurables	20 - 23, 26 - 31
Transportation and Communication	40 - 49
Wholesale Trade - Durables	50
Wholesale Trade - Nondurables	51
Retail Trade - Eating and Drinking Places	58
Retail Trade - Department, Apparel and Food Stores	53, 54, 56
Retail Trade - Other	52, 55, 57, 59
Finance, Insurance and Real Estate	60 - 69
Services - Hotels	70
Services - Personal and Business	72, 73
Services - Medical	80
Services - Educational	82
Services - Other	75 - 79, 81, 83 - 89
Public Sector	90 - 98

Appendix Table 2  
Sample Summary Statistics

	Mean	Std Dev
<b>Main Analysis Variables</b>		
percent of new hires from nonemployment [ $\beta_{ist}$ ]	70.196	10.674
log of percent of new hires from nonemployment [ $\ln(\beta_{ist})$ ]	4.240	0.155
new hires in the state-industry [ $m_{ist}$ ]	60.130	39.509
log of new hires in the state-industry [ $\ln(m_{ist})$ ]	3.867	0.737
<b>Unemployment and Vacancy Variables</b>		
state unemployment [ $U_{st}$ ]	8.018	2.198
log of state unemployment [ $\ln(U_{st})$ ]	2.043	0.283
local help-wanted rate [ $V_{st}$ ]	3.086	0.896
log of local help-wanted rate [ $\ln(V_{st})$ ]	1.079	0.324
<b>Variables Affecting Search and Matching [<math>A_{ist}</math>]</b>		
change in log industry employment <sup>a</sup>	0.062	5.092
log share of state employment	-2.721	0.807
potential weekly UI benefit <sup>b</sup>	23.484	7.839
<b>State Labor Force Demographics [<math>Z_{st}</math>]</b>		
proportion female	0.439	0.019
proportion under 25	0.257	0.015
proportion over 55	0.131	0.014
proportion white	0.795	0.076
proportion unionized	0.152	0.074
<b>Firm Demographics [<math>F_{ist}</math>]</b>		
log of average monthly employment	5.899	1.236
log of average per capita wage bill	1.258	0.651
<b>State Distribution [<math>D_s</math>]</b>		
proportion of observations from Georgia	0.357	0.479
proportion of observations from Louisiana	0.182	0.386
proportion of observations from Missouri	0.288	0.453
proportion of observations from South Carolina	0.174	0.379
<b>Total Number of Observations = 1060</b>		

<sup>a</sup>Multipled by 100.

<sup>b</sup>Deflated to constant dollars.

Appendix Figure 1  
Comparison of State and Local  
Labor Market Conditions

