New Microfoundations for the Aggregate Matching Function

Margaret Stevens*

December 19, 2005

Department of Economics, University of Oxford
margaret.stevens@economics.ox.ac.uk

Abstract

Empirical studies of the aggregate labor market matching function have favored a Cobb-Douglas functional form, for which there are no microfoundations in the existing literature. I present a new model for the matching process, based on a “telephone-line” Poisson queuing process, which, unlike other microeconomic approaches, can be integrated directly into standard theoretical search models. This implies a CES matching function, approximately Cobb-Douglas when search costs are approximately linear. The model allows empirical estimates of matching function parameters to be interpreted in terms of the costs and benefits of search.

Keywords: aggregate matching function, search, recruitment, search externalities

JEL classification: J31.

*I would like to thank seminar participants at Oxford University, and at IZA Berlin, for useful discussion of an earlier version of this paper. I am particularly grateful to Godfrey Keller, and to the Editor and two anonymous referees, for helpful and constructive comments.
1 Introduction

The matching function is the lynchpin of search and matching models of the labor market. The rate of matching, \( m \), is taken to be a function of the stock of unemployed workers, \( u \), and the stock of vacant posts, \( v \), in the market. If there were no frictions, matching would be instantaneous (and the number of matches would be determined by the short side of the market). But when workers and firms have to engage in a costly and time-consuming process of search to find each other, the matching function captures the technology that brings them together. It encapsulates search and matching frictions, allowing a more realistic description of the labor market, and of unemployment.

The matching function is by definition the source of frictional unemployment in such models; it may be the source of additional inefficiencies, due to search externalities, since each searching worker or firm affects the rate of matching for other agents. But the interpretation of both empirical and theoretical results is limited by the assumption that the matching function is an exogenous “black box”, unrelated to agents’ behavior or to other features of the labor market environment.

In this paper, I present simple microfoundations for the aggregate matching function. The model has several advantages: it can be directly integrated into standard search and matching models; it is consistent with, and guides the interpretation of, empirical evidence; and it provides insight into the conditions for efficiency in search models.

1.1 The search for microfoundations

The twin objectives in establishing microfoundations for the matching function are, first, to gain a better understanding of the nature of frictions and hence justify its use as a modeling device, and secondly to determine the form of the function \( m(u, v) \). There is general agreement on some desirable basic properties for this function: it should be increasing and concave in both arguments, with \( m(0, v) = m(u, 0) = 0 \) since matches can-
not occur unless there are agents on both sides of the market. Constant returns is often imposed, and arguably reasonable, although some models (for example, the well-known model of Diamond, 1982) use matching functions with increasing returns, resulting in a thin market externality and multiple equilibria. It may also be convenient theoretically to assume (as in Acemoglu and Shimer, 1999) that the rate of matching tends to infinity if either $u$ or $v$ does so; this helps to guarantee the existence of equilibrium, although the underlying justification is not obvious.

A variety of functional forms is suggested by different specifications of the matching process. One category of models is derived from static “urn-ball” processes: suppose that in each period a proportion $\alpha$ of unemployed workers each place a ball (job application) in a randomly chosen urn (a job vacancy), and the employer fills the vacancy by selecting a ball from the urn at random. Here, matching frictions result from a coordination problem between workers leading to congestion – some urns receive several balls and others none. In addition, the parameter $\alpha$ represents the search intensity of unemployed workers, which may be low if search is costly. The expected number of matches is equal to the number of urns with at least one ball:

$$m = v \left( 1 - \left( 1 - \frac{1}{v} \right)^{\alpha u} \right)$$

(1)

This matching function can be approximated for large $v$ (holding $\frac{u}{v}$ constant) by:

$$m = v \left( 1 - e^{-\alpha u/v} \right)$$

(2)

This is the standard urn-ball functional form, obtained by (among others) Hall (1979) and Peters (1984). For a continuous-time version we could suppose that each worker sends out applications at constant flow rate $\alpha$ to randomly chosen vacancies. Then the

\footnote{Petrongolo and Pissarides (2001) provide a comprehensive survey.}
expected number of vacancies receiving at least one application in a time period of length $dt$ is $v \left(1 - e^{-\alpha u dt/v}\right)$, as in the urn-ball model above. Letting $dt$ tend to zero gives a Poisson matching rate that is linear in unemployment:

$$m = \alpha u$$  \hfill (3)

Matching functions (1) to (3) represent an asymmetric view of the matching process, in which search is undertaken by workers, while firms are passive recipients of applications. Mortensen and Pissarides (1999) extend the continuous-time model to allow for simultaneous search on both sides of the market, so that firms separately make contact with workers at recruitment rate $\gamma$. This leads to a symmetric linear matching technology:

$$m = \alpha u + \gamma v$$  \hfill (4)

Despite the intuitive appeal of urn-ball models, the resulting functions do not have particularly desirable theoretical properties and cannot easily be integrated into standard search models, most of which are continuous-time models treating workers and firms symmetrically. The ubiquitous function (2), which does at least satisfy the basic requirements outlined at the beginning of this section, is derived from a static, asymmetric, model. In the continuous-time equivalent (3) the main attractive feature of the urn-ball model has been lost: the probability that any vacancy receives more than one application is of order $dt^2$ so the congestion effect disappears in the limit.\footnote{Blanchard and Diamond (1994) avoid this result, obtaining (2) in a continuous-time setting, by assuming that while workers apply for jobs at a constant flow rate, vacancies are posted for an exogenous discrete length of time. This ensures that congestion remains, but begs a question as to why vacancies do not close when they have received enough applications.} Furthermore, the linear technologies (3) and (4) do not behave well when either $u$ or $v$ is small relative to the
other, and in particular do not satisfy the basic requirement that \( m(0,v) = m(u,0) = 0 \).

An alternative approach attributes matching frictions to mismatch. It is assumed that agents match instantaneously within markets, but workers and firms may find themselves in a market where no suitable match is currently available for them.\(^3\) If there are many submarkets, and no mobility between them, a log-normal distribution of the vacancy-unemployment ratio across micromarkets implies a CES-type aggregate matching function (see Petrongolo and Pissarides, 2001). Shimer (2005) has Poisson-distributed numbers of workers and firms in each market, and slow movement between markets; in numerical calculations the aggregate matching function is indistinguishable from a Cobb-Douglas form.

Urn-ball and mismatch models describe the process of matching, but abstract from the heterogeneities or information problems that are usually assumed to be the source of frictions. Lagos (2000) aims to derive a matching function in which frictions arise endogenously. Agents choose between heterogeneous locations, knowing both the payoffs from trade and the expected number of buyers and sellers at each location. In equilibrium, some fail to match because too many of them choose locations where match payoffs are high. The aggregate matching function is \( m(u,v) = \min\{u, \phi v\} \) where \( \phi \) depends on the degree of heterogeneity between locations. Similarly in Burdett, Shi and Wright (2001) buyers choose between locations (in this case individual sellers), knowing their prices and capacities. Although prices are set strategically, and location choice is endogenous rather than random, the equilibrium matching function has the urn-ball form (1). In these two models matching frictions are in a sense endogenous, arising from the behavior of agents who trade off high match payoffs against the risk of not matching. However they can be traced to an underlying exogenous source, limited mobility: within a time

\(^3\)Stock-flow matching (Coles and Smith, 1998), which is sometimes interpreted as an alternative to the matching function approach, is also based on this assumption.
It is remarkable that, despite the apparent differences between existing approaches to establishing microfoundations for the matching function, the source of frictions is essentially the same in all of them: an implicit assumption of limited mobility, with an associated coordination problem. Agents do not coordinate in choosing where to search, and this matters because unmatched agents cannot costlessly and instantaneously change location. This is the essence of the matching problem in Lagos (2000) and Burdett, Shi and Wright (2001); similarly in urn-ball models, workers who choose the same urn as other workers cannot transfer to a different urn; and in mismatch models, workers who fail to match in one market cannot instantaneously transfer to another. In each case it is implicitly assumed that it takes time, or is otherwise costly, to transfer attention from one potential partner (or market) to another, and this mobility problem leads in turn to the possibility of congestion and coordination failure.

1.2 Empirical evidence

Surveying the empirical evidence, Petrongolo and Pissarides (2001) conclude that the urn-ball functional forms do not perform well. They summarize the wealth of support for a Cobb-Douglas matching function with constant returns to scale. Blanchard and Diamond (1990) for the US, and Burda and Wyplosz (1994) for four European countries, estimated a CES function; for all countries except Spain they obtained an elasticity of substitution between unemployment and vacancies not significantly different from unity, corresponding to the Cobb-Douglas case. Blanchard and Diamond obtained an estimate of around 0.4 for the unemployment elasticity and 0.6 for vacancies. Burda and Wyplosz, in contrast, found a higher elasticity for unemployment than for vacancies, although their results are not directly comparable with those for the US since they used unemployment outflow rather than new hires as the dependent variable. The lack of microfoundations,
however, makes it difficult to interpret empirical results – theory does not explain why
the elasticity of substitution should be unity, or why the unemployment elasticity should
be higher or lower than the vacancy elasticity, or higher in one market than another.

1.3 Theoretical considerations

A recurrent theoretical theme is the efficiency of search equilibrium. Since search exter-
nalities are a by-product of the matching function, efficiency considerations suggest some
further desirable properties for it. Hosios (1990) identified a general condition for social
efficiency: the elasticity of matching with respect to unemployment must be equal to
the worker’s share of the match surplus. More precisely, if the surplus is shared according
to a Nash bargain with shares $\beta$ and $1 - \beta$ for the worker and firm, then the decisions of
workers and firms to enter and leave the market, and to create and destroy matches, are
socially efficient if and only if first, the matching function has constant returns to scale
and secondly, the unemployment elasticity is equal to the worker’s bargaining share:

$$\frac{u}{m} m_u = \beta$$

Condition (5) is also sufficient for efficient search intensity choices, provided that two
further conditions are satisfied. First, intensities must be input-augmenting (Pissarides,
2000): that is, the matching function can be written as a function of $\alpha u$ and $\gamma v$, where
$\alpha$ and $\gamma$ are the search intensities of workers and firms. Secondly, the effect of search
intensity for individual agents must be proportional: that is, if the equilibrium intensity
is $\alpha$, the matching rate for a worker searching at intensity $\hat{\alpha}$ is given by $\frac{\hat{\alpha} \cdot m}{\alpha \cdot m}$.

The Hosios condition (5) is somewhat unsatisfactory: there is no reason why it should
be satisfied, and yet nothing to explain why it is not. Pissarides (2000, page 198) argues

---

4Efficiency corresponds to maximization of the aggregate surplus, subject to the matching frictions.
that we are not likely to find intuition for it, since it “relates a parameter of the resolution of bargaining conflict to a parameter of the technology of matchings”. There is little more to be said if the matching function is regarded as an exogenous black box: without microfoundations, we cannot say what determines the unemployment elasticity.

1.4 A model of time-consuming search

As shown in section 1.1, existing models of microfoundations rely on an assumption of limited mobility. The assumption that searching agents take time to transfer attention from one potential match to another is usually made without explicit justification. Nevertheless it is plausible: Moen and Rosen (2004) claim that frictions are due to “the costs and time delays associated with writing and processing applications, with identifying firms with vacancies, or with testing applicants.” Note that match heterogeneity must be an integral part of this story: such costs are significant when agents need to evaluate each potential match separately; they would not arise if all matches were the same ex post, or if agents could separate into sub-markets according to characteristics that were evident ex ante.

In what follows I present a model that makes the time-consuming nature of search explicit. I start from the presumption that matches are heterogeneous: the value of a match depends on characteristics of both parties that become evident only when they inspect each other. Hence agents search for a good match: that is, they consider potential matches one after another, taking time to evaluate each one, until they find one that is sufficiently attractive to both parties.

I model the allocation of time to search by both workers and firms in a continuous-time framework, which can be easily integrated into standard theoretical search models.

\footnote{An exception is the taxicab model of Lagos (2000), in which locations have a geographical interpretation so travel time is a natural requirement.}
The process is a symmetric continuous-time analogue of an urn-ball model, in which time-consuming search creates congestion and coordination failure. The resulting matching function is CES, and approximately Cobb-Douglas when marginal search costs are approximately constant. Hence it is consistent with existing empirical evidence, but also identifies the determinants of the unemployment elasticity, and of the elasticity of substitution between unemployment and vacancies, giving us an interpretation of empirical estimates of these parameters. Finally, since the unemployment elasticity is endogenous, it offers a different perspective on efficiency conditions in search models.

2 The Matching Model

The specification of the labor market is standard, of the type that has been used extensively to model frictional unemployment and to explore the implications of search externalities (see for example, Hosios, 1990, and Pissarides, 2000). There is a mass of ex ante identical workers, and a mass of ex ante identical firms, and each firm can employ one worker. Let \( u \) be the number of unemployed workers and \( v \) the number of vacant jobs. When an unemployed worker meets a firm with a vacancy, the flow productivity \( y \) of their match is realized, as a random variable from some continuous distribution. If the net surplus from the match is positive, the match is consummated and the surplus is shared according to a rule giving the worker an exogenous share \( \beta \in (0, 1) \).\(^6\) Productivity remains constant for the duration of the match, but matches are destroyed at exogenous rate \( \lambda \), in which case the worker re-enters unemployment, and the job becomes vacant.

All agents are risk neutral, and there is no discounting. Let \( Y_u \) and \( Y_v \) be the expected steady-state equilibrium income flows for an unemployed worker and a firm with a vacant

\(^6\)If wages are determined by ex post bargaining, \( \beta \) is the worker’s bargaining power. Alternatively \( \beta \) could be specified in a wage determination rule posted ex ante by the firm, as in directed search models (see, for example, Rogerson, Shimer and Wright, 2005).
job, respectively. Then the reservation productivity level is:

\[ z = Y_u + Y_v \]  

and the probability that a match with productivity \( y \) is acceptable is:

\[ p(z) \equiv \Pr[y > z] \]  

Since matches break down at rate \( \lambda \), the expected present value of the surplus of an acceptable match is:

\[ S(z) \equiv \frac{1}{\lambda} E[y - z | y > z] \]  

and the worker and firm obtain surpluses \( \beta S \) and \( (1 - \beta)S \) respectively.

2.1 The costs of search

Discovering the productivity of a potential match is time-consuming for both workers and firms. Unemployed workers spend time applying for jobs; we assume that the time required to produce an application is an exponentially-distributed random variable\(^7\) with mean normalized to 1. The worker allocates his time between search (preparing applications) and leisure; if he spends a proportion \( \alpha \) of each time period on search he obtains per-period utility \( b - C_w(\alpha) \), where \( b \) is constant and the utility cost of time spent on search, \( C_w \), is increasing and convex in \( \alpha \). Then, by the exponential assumption, he completes applications at Poisson rate \( \alpha \).

Recruitment is similarly costly for firms: an exponentially-distributed processing

\(^7\)Time required for search is assumed to be exponentially-distributed for both workers and firms, thus ensuring that matching rates and expected incomes are constant over time for each agent. This is convenient but not essential: for example with a zero discount rate a model with non-stochastic search times generates identical equilibrium conditions.
time, with mean $1/\gamma$, is required for each application considered by the firm. Here we can interpret $\gamma$ as the productivity of the recruitment manager, and the per-period cost of recruitment as his wage, $c + C_f(\gamma)$, which is increasing and convex in $\gamma$.

The parameters $\alpha$ and $\gamma$ thus represent the search and recruitment intensities of workers and firms. Initially we take them to be exogenous; later we allow agents to optimize.

### 2.2 The telephone-line matching technology

The model for the process by which the firms and workers meet is based on the classic “telephone-line” Poisson queueing process (Cox and Miller, 1965).

- Unemployed workers send applications (make calls) to firms at Poisson rate $\alpha$: that is, each application is completed in an exponentially-distributed time interval with mean $1/\alpha$, and then sent to a randomly-chosen vacancy.

- At any instant the firm is either waiting at the telephone to receive calls, or processing at Poisson rate $\gamma$. Specifically, after creating a vacancy, or answering a previous call, the firm spends time processing, and the time required is exponentially-distributed with mean $1/\gamma$.

- If a call arrives while the firm is busy processing, there is no answer; the worker fails to contact the firm, and begins a new application.

- When a call is answered the value of the potential match is revealed; if it is sufficiently productive the match is consummated; otherwise both parties continue to search.

This technology captures the same kind of congestion and coordination problem as in urn-ball models of matching. Since applications are not coordinated, and the firm must
give time to each one, some firms receive more than they can process while others are waiting for applications to arrive. By sending applications workers reduce the contact rate of other workers, and increase that of firms. Similarly processing by firms confers external benefits on workers and costs on other firms.

Some simplifying assumptions are built in to the process described above. The first is a timing assumption. It would be natural to suppose that the firm’s processing time is spent assessing an application after it is received; however to synchronize the search activity on both sides of the market it is convenient to assume instead that, like the worker, the firm prepares in advance to evaluate each potential match, and the productivity of the match is revealed at the instant that they meet. An alternative specification in which firms process applications after they arrive has similar properties, but there are further congestion effects: either the worker has to wait for a response before continuing to search, or the firm risks losing him to another firm before processing is complete.

The second assumption is that applications fail completely when a firm is busy; again, to relax this we would have to allow workers to wait, or to apply to several firms concurrently. This adds considerable complexity, with little additional insight. Our simplifications capture the essential feature that a period of time must be allocated by both parties to assess each potential match (causing congestion), while avoiding the need to keep track of queue lengths and multiple applications.

2.3 The matching rate with exogenous search intensity

Proposition 1 For the telephone-line matching technology, the aggregate matching rate in a steady-state equilibrium in which workers search at rate $\alpha$ and firms recruit at rate $\gamma$ is given by:

$$m(u, v, \alpha, \gamma, z) = p(z) \frac{\alpha u \gamma v}{\alpha u + \gamma v}$$

(9)
Proof: The total number of vacancies is \( v \). At any instant, a vacancy is in one of two states: it is either “waiting” for an application to arrive, or “processing” and hence unable to respond. Let \( v_0 \) be the stock of waiting vacancies. A processing vacancy becomes a waiting vacancy at rate \( \gamma \); hence the inflow to this stock is \( \gamma(v - v_0) \). The total number of applications sent out per unit time is \( \alpha u \), so the arrival rate of applications at a vacancy is \( \alpha u/v \), and the outflow from the stock of waiting vacancies is \( (\alpha u/v)v_0 \). Equating the flows and rearranging, we obtain the proportion of vacancies that are waiting in the steady state:

\[
\frac{v_0}{v} = \frac{\gamma v}{\alpha u + \gamma v}
\]

This expression is the probability that any individual application encounters a waiting vacancy and thus makes contact with a firm, so the aggregate contact rate is \( \alpha u v_0/v \). Multiplying by the probability that a contact results in a match, \( p \), gives the matching rate (9).

The matching function (9) is increasing and concave in \( u \) and \( v \), and has constant returns to scale.\(^8\) It is also increasing and concave in the search intensities \( \alpha \) and \( \gamma \), which are input-augmenting, and decreases with reservation productivity \( z \). The matching rate tends to zero as \( u \) or \( v \) tends to zero; as the number of vacancies tends to infinity the contact rate tends to \( \alpha u \), which is the rate of application,\(^9\) and as the number of unemployed workers becomes large, the contact rate approaches \( \gamma v \), the processing rate.

\(^8\)Berentsen, Rocheteau and Shi (2003) refer to (9) as the **additive matching-rate technology**, and note that with \( \alpha = \gamma \) this is the technology used in most monetary search models. In such models agents do not know whether other agents are buyers or sellers until they meet. So, for example, if there are \( v \) buyers and \( u \) sellers, and the sellers contact other agents randomly at a constant rate, the probability that the other agent is a buyer is \( \frac{v}{u + v} \), and the matching rate is proportional to \( \frac{uv}{u + v} \).

\(^9\)The linear technology (3) can be interpreted as the limit of the telephone-line technology as \( \gamma \to \infty \).
2.4 Equilibrium conditions

Proposition 1 shows that the telephone-line matching process generates an aggregate matching rate with desirable properties; we now show that when the process is embedded within the labor market model it yields standard labor market equilibrium conditions.

**Proposition 2** For the telephone-line matching technology, in a steady-state equilibrium with search intensities $\alpha$ and $\gamma$ the expected incomes for workers and firms are:

\[
Y_u = b - C_w(\alpha) + \frac{m}{u} \beta S \quad (11)
\]

\[
Y_v = -c - C_f(\gamma) + \frac{m}{v} (1 - \beta) S \quad (12)
\]

where $m$ is the aggregate matching rate (9). The matching rates for workers and firms searching at intensities $\hat{\alpha}$ and $\hat{\gamma}$ are given by:

\[
m_w(\hat{\alpha}) = \frac{\hat{\alpha} m}{\alpha u} \quad \text{and} \quad m_f(\hat{\gamma}) = \frac{\hat{\gamma} m}{\gamma v} \quad (13)
\]

and the individually-optimal search intensities satisfy:

\[
C'_w(\alpha) = \frac{m}{\alpha u} \beta S \quad \text{and} \quad C'_f(\gamma) = \frac{m}{\gamma v} (1 - \beta) S \quad (14)
\]

Note in particular the result that the individual matching rates (13) are proportional: the matching rate for an individual worker or firm is given by his own rate of search, multiplied by the number of matches per efficiency unit of search in the market. With a “black box” aggregate matching function this property is often assumed, as in Pissarides (2000). For the telephone-line matching technology, however, it is not an additional assumption, but follows from the specification of the matching process.

**Proof of Proposition 2:** Consider a worker sending out applications at current rate $\hat{\alpha}$, when his future rate, and the rate used by other workers at all times, is $\alpha$. The probability
that a job application encounters a “waiting” vacancy and hence results in a contact is given by the expression (10). So his matching rate is $m_w = \hat{\alpha}p \left( \frac{m}{u} \right) = \frac{\hat{\alpha}m}{u}$, and his expected income is

$$Y_u(\hat{\alpha}) = z - C_w(\hat{\alpha}) + m_w(\hat{\alpha})\beta S$$

(Note that since he reverts to rate $\alpha$ in future, he obtains the equilibrium surplus in the event of a match.) Maximising with respect to $\hat{\alpha}$ and setting $\hat{\alpha} = \alpha$ gives the first order condition for search intensity (14) and equilibrium income (11).

The derivation of the firm’s expected income is less straightforward, because a match can only occur when the vacancy is “waiting” for applications. The reservation income $Y_v$ is the expected income of a firm while it is processing, since this is the fall-back state if the firm decides not to enter a potential match. To derive $Y_v$, we will assume a positive discount rate $r$, and let $V, V_0$, and $J$ be the equilibrium expected present values of a processing vacancy, a waiting vacancy, and a filled job, respectively. Then we can write down Bellman equations for $V$ and $V_0$, and obtain $Y_v$ as the limit of $rV$ as $r$ tends to zero.

A vacancy finishes processing at rate $\gamma$, so the Bellman equation for $V$ is:

$$rV = -x - C_f(\gamma) + \gamma(V_0 - V)$$

A waiting vacancy receives applications at rate $\frac{\alpha u}{v}$, and an application results in a match with probability $p$. Hence:

$$rV_0 = -x - C_f(\gamma) + \frac{\alpha u}{v} \left( p(J - V_0) + (1 - p)(V - V_0) \right)$$

$$= -x - C_f(\gamma) + \frac{\alpha u}{v} \left( p(J - V) + (V - V_0) \right)$$

$$= -x - C_f(\gamma) + \frac{\alpha u}{v} \left( p(1 - \beta)S + (V - V_0) \right)$$

Subtracting these two equations and letting $r$ tend to zero we obtain:

$$V_0 - V = \frac{\alpha u}{\alpha u + \gamma v} p(1 - \beta)S = \frac{m}{\gamma v} (1 - \beta)S$$
For a firm currently processing at rate $\hat{\gamma}$, with future rate $\gamma$, expected income is:

$$Y_v(\hat{\gamma}) = -x - C_f(\hat{\gamma}) + \hat{\gamma}(V_0 - V)$$

$$= -x - C_f(\hat{\gamma}) + \frac{\hat{\gamma}m}{\gamma v}(1 - \beta)S$$

Thus $m_f = \frac{\hat{\gamma}m}{\gamma v}$ is the effective matching rate, allowing for the two states. As before, maximising with respect to $\hat{\gamma}$ and setting $\hat{\gamma} = \gamma$ gives the required equilibrium conditions.

To complete the determination of equilibrium we need to specify entry conditions to the market for both firms and workers, and a steady-state condition for employment.\(^\text{10}\)

These conditions, together with (6), (9), (11), (12), and (14), determine the values of $u$, $v$, $\alpha$, $\gamma$ and $z$.

### 2.5 The matching rate with endogenous search intensity

The matching function (9) determines the matching rate $m$ conditional upon search intensities $\alpha$ and $\gamma$. Equations (14) determine (implicitly) the worker’s and firm’s optimally chosen search and recruitment intensities, $\alpha = \alpha^*(u,v,z)$ and $\gamma = \gamma^*(u,v,z)$.

By substituting these back into the conditional matching function, we can obtain the matching function with endogenous search, or unconditional matching function:

$$m^*(u,v,z) = m(u,v,\alpha^*(u,v,z),\gamma^*(u,v,z),z)$$

The function $m^*$ is of empirical interest, since search and recruitment effort are not normally observed. The relationship between the two functions $m$ and $m^*$ depends on

\(^{10}\)Typically in a search model, it is assumed that there is an exogenous mass $n$ of workers, giving a steady-state condition $m = \lambda(n - u)$; and free entry of firms, so that $Y_v = 0$. However, our analysis is equally valid in a model with a fixed mass of firms as well as workers, or with imperfectly elastic supplies of agents to the market.
the form of search and recruitment costs. If the cost functions have constant and identical elasticities, we can derive the unconditional matching function explicitly:

**Proposition 3** With the telephone-line matching technology and search cost functions:

\[ C_w(\alpha) = \frac{c_w}{k} \alpha^k \quad \text{and} \quad C_f(\gamma) = \frac{c_f}{k} \gamma^k \]  

(16)

the unconditional matching function is:

\[ m^*(u, v, z) = m_0(z) \left( \frac{1}{\bar{\eta}^{1-\rho} u^{-\rho} + (1 - \bar{\eta})^{1-\rho} v^{-\rho}} \right)^{\frac{1}{\rho}} \]  

(17)

where

\[ \rho \equiv 1 - \frac{1}{k}, \quad \bar{\eta} \equiv \frac{(1 - \beta)/c_f}{(1 - \beta)/c_f + \beta/c_w}, \quad m_0(z) \equiv p(z) \left( \frac{p(z) S(z)}{c_w/\beta + c_f/(1 - \beta)} \right)^{\frac{1-\beta}{\rho}} \]  

(18)

**Proof:** With the cost functions (16), the first-order conditions (14) give \( c_w \alpha^k = \frac{m \beta S}{u} \) and \( c_f \gamma^k = \frac{m (1 - \beta) S}{v} \). Substituting for \( \alpha \) and \( \gamma \) in (9) and solving for \( m \) gives (17).

Note that we can also eliminate the search intensities from the equilibrium conditions (11) and (12) to obtain:

\[ Y_u = b + \frac{m^*}{u} \beta \rho S \]  

(19)

\[ Y_v = -c + \frac{m^*}{v} (1 - \beta) \rho S \]  

(20)

These equations, together with entry and steady-state conditions, describe the equilibrium in a transformed model in which we can effectively ignore the determination of search intensities: they are captured automatically if we use the unconditional matching function \( m^* \), and reduce the effective surplus from \( S \) to \( \rho S \) to allow for expenditure on search.
3 The Properties of the Matching Function $m^*$

Section 2 shows how the telephone-line matching technology provides a micro-model for the matching process that can be fully integrated into a standard search and matching model. Furthermore, it generates an aggregate matching function that has the properties conventionally required of black-box matching functions. Both the conditional and unconditional matching functions $m$ and $m^*$ have the desirable basic properties identified in section 1.1, including constant returns to scale.

The conditional function $m(u, v, \alpha, \gamma, z)$ has a specific form (CES in $u$ and $v$ with elasticity of substitution between unemployment and vacancies equal to a half), but more interestingly for empirical purposes the unconditional matching function $m^*(u, v, z)$ defined by (17) has a flexible CES form with parameters $\rho$ and $\bar{\eta}$. According to the model these parameters are in turn determined by the bargaining share $\beta$ and the parameters of the cost functions. Several nice properties follow immediately:

1. The elasticity of substitution between unemployment and vacancies is determined by the elasticity of search and recruitment costs.

For the CES function (17), the elasticity of substitution between $u$ and $v$ is given by: $\sigma^* = 1/(1 + \rho)$. Hence from (18) it depends on the elasticity of the search cost functions, $k$:

$$\sigma^* = \frac{1}{2 - 1/k}$$

Thus the elasticity of substitution is inversely related to the cost elasticity $k$. When $k$ is very high, $\sigma^*$ is close to a half, which is the elasticity of substitution of the conditional matching function $m$. In this case search and recruitment effort are almost exogenous; in the limit both $\alpha$ and $\gamma$ are equal to one, and the conditional
and unconditional matching functions are identical:

\[ m^* = m = p(z) \frac{uv}{u + v} \]

As \( k \) decreases towards one, the cost function becomes more linear, and the elasticity of substitution between unemployment and vacancies increases towards one. Intuitively, if marginal costs are approximately constant, workers and firms can adjust their search intensities easily in response to changes in \( u \) and \( v \).

2. **The elasticity of matching with respect to unemployment is determined by the rate of return to search for firms relative to workers.**

Log-differentiating (17) gives the unemployment elasticity of the unconditional matching function:

\[ \eta^* \equiv \frac{\partial \log m^*}{\partial \log u} = \frac{\bar{\eta}^{1-\rho}u^{-\rho}}{(1-\bar{\eta})^{1-\rho}u^{-\rho} + (1-\bar{\eta})^{1-\rho}v^{-\rho}} \]  

\( \eta^* \) is increasing in \( \bar{\eta} \), which measures (see (18)) the rate of return to recruitment for firms, \( (1 - \beta)S/c_f \), relative to the rate of return to search for workers, \( \beta S/c_w \). Hence, the unemployment elasticity is high when the worker’s rate of return to search is low relative to that of firms.

3. **In equilibrium, the unemployment elasticity \( \eta^* \) is equal to the proportion of search activity undertaken by firms.**

It can be verified directly from (17), using the first-order conditions for \( \alpha \) and \( \gamma \), that the unemployment elasticity of \( m^* \) satisfies:

\[ \eta^* = \frac{\gamma v}{\alpha u + \gamma v} \]

This expression gives us an intuitive interpretation of the elasticity: from (10)
above, it is the probability that an individual application makes successful contact with a firm, or equivalently the proportion of total search effort undertaken by firms. Thus, a high unemployment elasticity corresponds to a low congestion problem facing workers. Symmetrically, the elasticity with respect to vacancies is equal to the probability of contact per unit of recruitment effort, which is the proportion of search effort undertaken by workers.

4. In equilibrium, the unemployment elasticity $\eta^*$ of the unconditional matching function is equal to the unemployment elasticity $\eta$ of the conditional matching function.$^{11}$

From equation (9) the unemployment elasticity of the conditional matching function $m(u, v, \alpha, \gamma, z)$ is given by

$$\eta \equiv \frac{\partial \log m}{\partial \log u} = \frac{\gamma v}{\alpha u + \gamma v}$$

5. The matching rate decreases with reservation productivity $z$, and increases with the expected surplus from a contact, $p_S$.

Recall that for the conditional matching function $m$, the matching rate is proportional to the probability $p$ of an acceptable match, which decreases with $z$. Similarly for the unconditional matching function $m^*$ the reservation productivity affects the matching rate through the scale term $m_0(z)$, which depends on both the probability $p$ of matching, and the expected surplus from a contact, $pS$, which is also decreasing in $z$. A higher surplus (holding $z$ constant) increases the investment in search on both sides of the market and thus raises the matching rate.

6. When the cost elasticity $k$ is close to one, the unconditional matching function $m^*$

$^{11}$We show in section 4 below that the equality of the elasticities here is a special case of a more general result.
is approximately Cobb-Douglas with constant unemployment elasticity equal to $\bar{\eta}$.

When $k$ is close to one (the cost function is almost linear), $\rho$ is close to zero and the elasticity of substitution $\eta^*$ between unemployment and vacancies is approximately one. Hence the unconditional matching function (17) is close to its Cobb-Douglas limiting form. From the expression above for the elasticity we obtain:

$$\eta^* \approx \bar{\eta} \equiv \frac{(1 - \beta)/c_f}{(1 - \beta)/c_f + \beta/c_w}$$

It is, indeed, intuitive that the unemployment elasticity is approximately constant in the case when search activity can easily respond to market conditions. When there are many more workers than firms, for example, the elasticity of matching with respect to unemployment will tend to be low, for given search effort. But search effort on both sides of the market responds to this asymmetry: workers search less, reducing congestion, and firms search more, raising the matching rate of workers and offsetting the asymmetry.

Note, however, that the function $m^*$ does not have a well-defined limit as $\rho \to 0$ (that is, as $k \to 1$) unless $pS = c_w/\beta + c_f/(1 - \beta)$. With linear costs, this condition is necessary for the first-order conditions for search intensities to have an interior solution.

7. When the number of unemployed workers becomes large relative to vacancies, the matching rate is determined by the recruitment rate; similarly when the number of vacancies is large it is determined by the application rate.

Again using (17) and the first-order conditions, it can be verified that in equilibrium:

$$\lim_{u \to \infty} m^*(u, v, z) = p\gamma v$$

Thus, as for the conditional function, the contact rate is equal in the limit to the
recruitment rate $\gamma v$. In particular, note that the matching rate is finite, although when $k$ is close to one it is high because firms choose high intensity $\gamma$. It is only the limiting Cobb-Douglas matching function that has the property that the matching rate tends to infinity with $u$ or $v$.

3.1 Empirical implications

These results provide an interpretation of existing empirical estimates. Blanchard and Diamond (1990) estimated a CES aggregate matching function on US data, obtaining an estimate of the elasticity of substitution between unemployment and vacancies of $\sigma^* = 0.74$. In the telephone-line model, this corresponds to a plausible value for the cost elasticity of $k = 1.54$. Their estimate of $\sigma^*$ was not significantly different from 1, corresponding to the Cobb-Douglas limiting case. Note, however, that our model restricts $\sigma^*$ to lie in the range 0.5 to 1.0, and it would also be interesting to test the hypothesis $\sigma^* = 0.5$, representing the case of exogenous search effort.

Blanchard and Diamond’s corresponding estimate for the unemployment elasticity was 0.48. In all their specifications they obtained a lower value for the unemployment elasticity than for the vacancy elasticity, as did Anderson and Burgess (2000), using US state-level data and a Cobb-Douglas specification. The interpretation provided by our model is that the rate of return to search is higher for workers than for firms, and a higher proportion of total search effort is therefore undertaken, in equilibrium, by workers. Furthermore, by property 4, although empirical studies estimate the unconditional matching function – they do not control for search effort – we can interpret their estimates of the unemployment elasticity as referring also to the underlying conditional matching function.

Where it is possible to estimate more disaggregated matching functions our model provides an interpretation of differences in coefficients between markets. Fahr and Sunde
(2001) estimate a log-linear (Cobb-Douglas) model disaggregated by occupation, age, and education. Unemployment elasticities are high (0.57 to 0.6) for technical, craft, service and white collar occupations, compared with 0.46 for low-skilled occupations. The interpretation would be that for the low-skilled occupations the rate of return to search for firms is relatively low and hence that a higher proportion of search is undertaken by workers. Fahr and Sunde also obtain (on average across occupations) higher constant terms for more educated workers. From Property 5 this is consistent with higher surpluses (relative to search costs) in markets for educated workers, giving a greater incentive to search intensively.

The model draws attention to an empirical problem: the dependence of the matching rate on the reservation productivity $z$. If it is not possible to control for $z$, or equivalently for the probability that a contact results in a match, estimates of the elasticity of matching with respect to unemployment and vacancies may be biased, depending on the market entry conditions which determine $z$. For example, with a fixed mass of workers and free entry of firms, the equilibrium value of $z$ depends on market tightness $v/u$. The problem is not specific to the unconditional matching function derived here: it would arise even if we could control for search intensity, in any setting where the matching rate differs endogenously from the contact rate. If, as argued in section 1, the source of matching frictions is the need for agents to inspect each other before deciding whether to match, it is a potential problem for any attempt to estimate matching functions. The issue is most serious for aggregate estimates; for disaggregated matching functions the reservation productivity may be determined by mobility between markets and hence exogenous for individual markets.\footnote{\textsuperscript{12}I am grateful to a referee for this point.}

Finally, it should be noted that there are other potential sources of bias, depending on the data used. For example, heterogeneous agents may make different choices of
search intensity. On-the-job search may also be important: Blanchard and Diamond (1990) estimate that 15% of new hires come directly from employment. The empirical applicability of the model would be enhanced if it were extended to allow employed workers to search, perhaps by giving them an endowment of time to be allocated between search and leisure, smaller than the time available to unemployed workers.

3.2 Efficiency

Since the conditional matching function \( m \) has constant returns and input-augmenting search, the equilibrium is efficient if and only if the Hosios condition holds. From the properties above we can see that the consistency between the underlying model and the transformed model based on the unconditional matching function \( m^* \) extends to the efficiency conditions: \( m^* \) has constant returns and from Property 4 the condition that the unemployment elasticity must be equal to the worker’s bargaining share can be applied to \( m^*(u, v, z) \).

But it is also clear that the unemployment elasticity should not be regarded as an exogenous parameter. When the matching function is approximately Cobb-Douglas, for example, the unemployment elasticity is close to \( \bar{\eta} \), which depends (see (18)) on the bargaining share and the cost parameters. This endogeneity does not mean that there is any tendency towards efficiency: if, for example, the bargaining share \( \beta \) is high, then \textit{ceteris paribus} the unemployment elasticity will be low, because workers search more intensively than firms. So if \( \beta \) is high, efficiency requires that the costs of search for workers are also high, to moderate their search activity and prevent the unemployment elasticity from falling. What we have shown by establishing microfoundations is that the Hosios condition defines a relationship between the costs and benefits of search under which private incentives are optimal.
4 A Generalization

The results obtained in the previous section are for a particular matching technology that is of interest because it gives us a CES unconditional matching function. However the technique employed to obtain them is applicable more widely: we can allow for endogenous search costs in a matching model by a simple transformation from the conditional to the unconditional matching function. This is particularly useful empirically, because we do not normally observe search effort; in theoretical models it means that we can ignore search effort unless it is of specific interest. So we now derive some more general results, which hold for any constant elasticity cost functions, and any matching technology with standard properties. Specifically, we make the following assumptions in this section:

A1. The matching technology is increasing, concave, and homogeneous of degree 1 in $u$ and $v$, with input-augmenting search:

$$m = m(\alpha u, \gamma v)$$

(22)

and individual matching rates are proportional (that is, satisfy (13)).

A2. The cost functions for search and recruitment intensity have constant elasticities:

$$C_w(\alpha) = \frac{c_w}{k_w} \alpha^{k_w} \quad \text{and} \quad C_f(\gamma) = \frac{c_f}{k_f} \gamma^{k_f} \quad \text{where} \quad k_w, k_f > 1 \text{ and } c_w, c_f > 0$$

Under these assumptions equations (11), (12) and (14) can be written down immediately (by the usual “black box” approach). Substituting from the first-order conditions for

---

13 As before, the matching rate is proportional to the probability $p(z)$ of an acceptable match, but to reduce notation the dependence on $z$ is suppressed.
search intensities (14) into the general unconditional matching function (22), and writing:

\[
\rho_w \equiv 1 - \frac{1}{k_w} \quad \text{and} \quad \rho_f \equiv 1 - \frac{1}{k_f}
\]

we obtain an equation for the equilibrium matching rate \( m^* \):

\[
m^* = m \left( (\bar{S}m^*)^{1-\rho_w} \frac{u^{\rho_w}}{\bar{\eta}^{1-\rho_w}}, (\bar{S}m^*)^{1-\rho_f} \frac{v^{\rho_f}}{(1-\bar{\eta})^{1-\rho_f}} \right) \quad (23)
\]

where:

\[
\bar{\eta} \equiv \frac{(1 - \beta)/c_f}{(1 - \beta)/c_f + \beta/c_w} \quad \text{and} \quad \bar{S} \equiv \frac{S}{c_w/\beta + c_f/(1 - \beta)}
\]

Equation (23) defines \( m^* \) as an implicit function of \( u \) and \( v \) – the unconditional matching function. Then we have the following results for the function \( m^* \):

**Proposition 4** When the matching technology and search cost functions satisfy A1 and A2:

(i) The unconditional matching function \( m^* \) is increasing, concave, and homogeneous of degree 1 in \( u \) and \( v \).

(ii) The unemployment elasticity \( \eta^* \) of the unconditional matching function is related to the unemployment elasticity \( \eta \) of the conditional matching function by:

\[
\eta^* = \frac{\rho_w \eta}{\rho_w \eta + \rho_f (1 - \eta)}
\]

(iii) The Bellman equations for unemployment and vacancies can be written:

\[
\begin{align*}
Y_u &= b + \frac{m^*(u, v) \beta^* \rho^* S}{u} \\
Y_v &= -c + \frac{m^*(u, v) (1 - \beta^*) \rho^* S}{v}
\end{align*}
\]
where
\[
\beta^* \equiv \frac{\rho_w \beta}{\rho_w \beta + \rho_f (1 - \beta)} \quad \text{and} \quad \rho^* \equiv \rho_w \beta + \rho_f (1 - \beta)
\]

(iv) When the cost elasticities are equal \((\rho_w = \rho_f = \rho)\), the unconditional matching function is given explicitly by:

\[
m^*(u, v) = \bar{S} \frac{1 - \eta}{\eta} m(\bar{\eta}^{\rho - 1} u^{\rho}, (1 - \bar{\eta})^{\rho - 1} v^{\rho})^{\frac{1}{\rho}}
\]

the unemployment elasticities of \(m^*\) and \(m\) are equal:

\[
\eta^* = \eta
\]

and the elasticities of substitution are related by:

\[
\frac{1}{\sigma^*} - 1 = \rho \left( \frac{1}{\sigma} - 1 \right)
\]

PROOF: See Appendix.

The implication of Proposition 4 is that for any standard matching technology, provided we assume that search and recruitment intensities are input-augmenting and proportional with constant cost elasticities, we can ignore the search intensity choice problem and work with a transformed model based on the unconditional matching function \(m^*\), and a reduced surplus \(\rho^* S\).

In the case where the cost elasticities for workers and firms are equal, we can see that some of the results obtained in section 3 apply more generally. First, the unemployment elasticity of the underlying conditional function is the same as that of the unconditional function, so can be estimated without data on search intensities. Second, the effect of endogenous search intensity is to increase the elasticity of substitution between unem-
ployment and vacancies; and the lower is the cost elasticity \( k \), the more will workers and firms adjust their search in responses to changes in \( u \) and \( v \), and hence the higher is the elasticity of substitution between \( u \) and \( v \).

When the cost elasticities differ, there are two advantages for the party with higher cost elasticity. If, for example, the search cost elasticity for workers, \( k_w \), is high, then \( \rho_w \) is high. Then, first, the rate of matching \( m^* \) is more elastic with respect to the stock of unemployed workers, and secondly, workers have higher effective bargaining power \( \beta^* \).

Intuitively, when they are less able to adjust their search activity, they earn higher rents in equilibrium. Finally, with differing cost elasticities a transformed version of the Hosios condition applies in the transformed model: from result (ii) above \( \beta = \eta \) is equivalent to \( \beta^* = \eta^* \).

**Appendix**

**Proof of Proposition 4:**

(i) Using the homogeneity of \( m \), (23) can be written:

\[
\frac{m^*}{u} = m \left( \left( \frac{\bar{S}m^*}{\eta u} \right)^{1-\rho_w}, \left( \frac{\bar{S}m^*}{\eta u} \right)^{1-\rho_f} \left( \frac{v}{u} \right)^{\rho_f} \right)
\]

Hence \( m^*/u \) is an implicit function of \( v/u \), and homogeneity follows immediately. Now write (23) as:

\[
m^*(u, v) = m(X, Y)
\]

where \( X(m^*, u) \equiv (\bar{S}m^*)^{1-\rho_w} \bar{\eta}^{\rho_w} u^{\rho_w} \) and \( Y(m^*, v) \equiv (\bar{S}m^*)^{1-\rho_f} \bar{\eta}^{\rho_f} v^{\rho_f} \)

Log-differentiating:

\[
\frac{\partial \log m^*}{\partial \log u} = \left( (1 - \rho_w) \frac{\partial \log m^*}{\partial \log u} + \rho_w \right) \frac{\partial \log m}{\partial \log X} + (1 - \rho_f) \frac{\partial \log m^*}{\partial \log u} \frac{\partial \log m}{\partial \log Y}
\]

\[\Rightarrow \eta^* = ((1 - \rho_w)\eta^* + \rho_w) \eta + (1 - \rho_f)\eta^*(1 - \eta)\]
\[ \eta^* = \frac{\rho_w \eta}{\rho_w \eta + \rho_f (1 - \eta)} \]  

(24)

Hence \( m^* \) is an increasing function of \( u \), and similarly of \( v \).

Now differentiate again with respect to \( \log u \):

\[ \frac{\partial \eta^*}{\partial \log u} = \frac{\rho_w \rho_f}{(\rho_w \eta + \rho_f (1 - \eta))^2} \left( (1 - \rho_w) \eta^* + \rho_w \right) \frac{\partial \eta}{\partial \log X} + (1 - \rho_f) \eta^* \frac{\partial \eta}{\partial \log Y} \]

Homogeneity of \( m \) implies that \( \frac{\partial \eta}{\partial \log X} = -\frac{\partial \eta}{\partial \log Y} \) so:

\[ \frac{\partial \eta^*}{\partial \log u} = \frac{\rho_w \rho_f}{(\rho_w \eta + \rho_f (1 - \eta))^2} (\rho_w (1 - \eta^*) + \rho_f \eta^*) \frac{\partial \eta}{\partial \log X} \]

Also:

\[ \frac{\partial \eta^*}{\partial \log u} = \frac{u^2}{m^*} m_{11}^* \eta^* (1 - \eta^*) \] and similarly \( \frac{\partial \eta}{\partial \log X} = \frac{X^2}{m} m_{11} + \eta (1 - \eta) \)

Substituting both of these expressions in the previous equation:

\[ \frac{u^2}{m^*} m_{11}^* \eta^* (1 - \eta^*) \left( \frac{X^2}{m} m_{11} + \eta (1 - \eta) \right) = \left( \rho_w (1 - \eta^*) + \rho_f \eta^* \right) - \eta^* (1 - \eta^*) \]

Since \( m^* \) is a linear homogeneous function, \( m_{11}^* \leq 0 \) is necessary and sufficient for concavity. Since \( m \) is concave, both terms on the right-hand side are negative and \( m^* \) is concave.

(ii) was proved at (24) above.

(iii) Using the first-order conditions (14) to substitute for search intensities in the Bellman equations (11), (12) we obtain:

\[
\begin{align*}
Y_u & = b + \frac{m^*(u,v)}{u} \rho_w \beta S \\
Y_v & = -c + \frac{m^*(u,v)}{v} \rho_f (1 - \beta) S
\end{align*}
\]
and defining $\beta^*$ and $\rho^*$ gives the required result.

(iv) Putting $\rho_w = \rho_f = \rho$ in (23) and using the homogeneity of $m$ gives the expression for the unconditional matching function immediately. Putting $\rho_w = \rho_f = \rho$ in (24) gives $\eta^* = \eta$.

The elasticities of substitution are:

$$\sigma^* \equiv \frac{\partial \log(v/u)}{\partial \log(m_1^*/m_2^*)} \quad \text{and} \quad \sigma \equiv \frac{\partial \log(Y/X)}{\partial \log(m_1/m_2)}$$

Since the unemployment elasticities are equal:

$$\frac{\eta^*}{1-\eta^*} = \frac{\eta}{1-\eta}$$

$$\Rightarrow \frac{um_1^*}{vm_2^*} = \frac{Xm_1}{Ym_2}$$

$$\Rightarrow \log \left( \frac{m_1^*}{m_2^*} \right) - \log \left( \frac{v}{u} \right) = \log \left( \frac{m_1}{m_2} \right) - \log \left( \frac{X}{Y} \right)$$

and differentiating with respect to $\log(v/u)$:

$$\frac{1}{\sigma^*} - 1 = \left( \frac{1}{\sigma} - 1 \right) \frac{\partial \log(Y/X)}{\partial \log(v/u)}$$

$$\Rightarrow \frac{1}{\sigma^*} - 1 = \rho \left( \frac{1}{\sigma} - 1 \right) \quad \blacksquare$$

References


