

Matching, Search, and Heterogeneity*

Robert Shimer
Princeton
shimer@princeton.edu

Lones Smith
Michigan
lones@umich.edu

July 2000

Abstract

This paper explores optimal search and matching behavior in economies with heterogeneous agents. Optimal policies may be nonstationary since cyclical behavior can achieve outcomes that are unattainable in steady state or can reduce the cost of attaining a given outcomes. The paper therefore characterizes the social optimum without imposing any stationarity restrictions. It then contrasts optimal policies with equilibrium outcomes by finding search taxes that decentralize the social optimum. The optimal tax schedule implies that high productivity agents generally search too little and match too frequently, while low productivity agents do the opposite. We interpret these results in terms of thick-market and congestion externalities.

*This is a major revision of material contained in an early version of “Assortative Matching and Search” and in “Normative Implications of Heterogeneity in Search”

1 Introduction

Consider the following canonical economic situation: a pair of impatient agents meet and must decide whether to match. The benefit is that they prefer to be matched rather than to be unmatched. This is weighed against an opportunity cost, as matching precludes further search for more suitable partners. Examples include the decision of a man and a woman to marry; the decision of a home buyer and home seller to sign a sales contract; and the decision of a worker and firm to enter into an employment relationship.

Economists have long recognized that the decision of a pair of agents to match involves an externality, since it alters the meeting opportunities of other unmatched agents. In general, two offsetting effects have been identified: when there are more agents looking for partners, it is easier to locate a potential partner (the thick market externality); and when there are more agents looking for partners, competing searchers more rapidly locate and match with potential partners (the congestion externality). Mortensen (1982) and Hosios (1990) show that the correct allocation of property rights may ensure optimal search behavior under special circumstances. However, these papers assume that agents are *ex ante* identical. Although this permits them to examine optimal search behavior, they must sidestep the matching dimension of the problem.

This paper studies an environment with nontrivial search and matching decisions. Heterogeneous unmatched agents choose the intensity of their search activities; and when they meet other unmatched agents, they must decide whether to match. Both decisions impose externalities, although in the model considered, these cancel out if agents are homogeneous; the decentralized equilibrium is efficient. In contrast, when agents are heterogeneous, the decentralized equilibrium is generically inefficient.

Almost all existing work on heterogeneous-agent search and matching models focuses on steady state behavior.¹ This is a reasonable restriction for equilibrium analysis, since Shimer and Smith (2000) show that steady state equilibria exist under fairly general conditions. If the economy — i.e. the state variables and the strategies of all other agents — is in steady state, any individual agent should use a stationary strategy. But that logic does not carry over to our analysis of social

¹To our knowledge, the only exception is Burdett and Coles (1998), which explores nonstationary equilibria in a nontransferable utility search model. Sattinger (1995) characterizes socially optimal search behavior in a transferable utility model, but restricts attention to steady states.

optima. A social planner does not treat the strategies of other agents as given, and so can always use a nonstationary policy.

To show that this is not merely a technical concern, we develop two simple examples that illustrates the optimality of limit cycles in generic models. In one example, limit cycles arise from the possibility of repeatedly reaching points in the state space that cannot be sustained in the steady state. This logic is inherent in all search models with nontrivial matching decisions. In the other model, the cycles arise from the possibility of using time-varying search intensities to coordinate search activities. Because of these examples, we are forced to pursue a nonstationary characterization of optimal search and matching behavior.

A direct comparison of behavior in the decentralized equilibrium and social optimum does not yield a precise characterization of the nature of the inefficiency, as it is convoluted by the complex and changing interaction between agents. Instead, we characterize the inefficiency by describing a ‘Pigouvian’ tax scheme that decentralizes the social optimum.² One such scheme is a tax on search activity for the least productive agents, and subsidy for the most. The tax/subsidy has two effects: agents who receive a subsidy increase the intensity of their search activity; and they are more reluctant to accept matches, since search is less costly. From this we conclude that in a decentralized equilibrium, the most productive agents do not search hard enough and match too frequently. The opposite holds for the least productive.

We can interpret this finding in terms of the thick-market and congestion externalities. When an agent searches harder or decides to remain unmatched, she makes it easier for other agents to match with her (the thick-market externality) but makes it harder for other agents to meet each other (the congestion externality). If she is more productive than average, the thick-market externality is more important; the ease of meeting her outweighs the difficulty of meeting others. But if she is less productive than average, the congestion externality is more important. It follows that previous studies of efficiency in search equilibrium hinge on the homogeneity assumption. As long as the decision of an agent to match or to search harder alters the distribution of meetings for other agents, the decentralized equilibrium will be inefficient.³

²This is not a public finance exercise, and we so we make no effort to discuss whether these taxes could be implemented in practice.

³Models of ‘directed search’, in which agents can choose whom to meet, may still yield an efficient equilibrium. The key property is that if one agent matches, it is no more difficult to

Section 2 develops a model of search. In Section 3, we develop two simple examples to show that optimal search behavior need not be stationary. We solve for the social optimum in Section 4 and the decentralized equilibrium in Section 5. Then in Section 6, we compare these allocations using Pigouvian taxes. Since the search process may be crucial to our results, we discuss our specification and some alternatives in Section 7.

2 Model

We develop a continuous time model in which a measure 1 continuum of agents search for partners. All agents are risk-neutral, infinitely-lived, and discount the future at rate $r \geq 0$. Agents are identified by their type $i \in \{1, \dots, N\}$, with a fraction $\ell_i > 0$ of agents being of type i . At any point in time t , each agent may be either matched or unmatched. Let $m_{ij}(t)$ denote the measure of type i agents matched with type j agents at time t . This is also the measure of (i, j) matches for $i \neq j$, but $m_{ii}(t)$ is *twice* the measure of (i, i) matches. Then $u_i(t) = \ell_i - \sum_{j=1}^N m_{ij}(t)$ denotes the measure of unmatched type i agents.

Each unmatched agent i searches for partners with endogenous intensity $\rho_i(t)$ at cost $c(\rho_i(t))$.⁴ Search frictions capture the following story: Unmatched agents meet in crowded locations ('bars'). An agent's search intensity determines how frequently she visits a bar. When she goes to one, she is guaranteed to meet one other agent, randomly selected from those at the bar. As a result, the rate at which she meets another unmatched agent depends only on her own search intensity. However, the potential partner is drawn randomly from the unmatched population, with likelihood proportionate to his contemporaneous search intensity. Thus an agent who searches with intensity ρ meets a type j agent at time t with instantaneous probability $\rho \rho_j(t) u_j(t) / \sum_{k=1}^N \rho_k(t) u_k(t)$. Note that the aggregate matching rate is linear in the unmatched rate, $\sum_{i=1}^N \rho_i(t) u_i(t)$. In particular, this generalizes Pissarides's (2000) matching function with endogenous search intensity (Chapter 5) to an environment with heterogeneous agents.

When two agents i and j meet, the social planner must decide whether they

contact other similar agents (Acemoglu and Shimer 1999, Shi 1998).

⁴Since we do not impose any restrictions on the search cost function, exogenous search intensity is a special case of this model. Let $c(\rho) = 0$ if $\rho = \bar{\rho}$, and $c(\rho) = \bar{c} \gg 0$ otherwise. Then search intensity is 'endogenously' equal to $\bar{\rho}$.

should match. If they match, they produce a flow output $f_{ij} \equiv f_{ji}$, the only extrinsic source of heterogeneity in this model. However, matched agents forego the opportunity to search, and so never meet any unmatched agents. At any time, the social planner may endogenously terminate a match. In addition, all matches are exogenously terminated according to a Poisson process with arrival rate $\delta > 0$. After either an endogenous or exogenous termination, the pair starts off again unmatched. Following the literature, the objective of the social planner is utilitarian, which here amounts to simply maximizing the present value of output net of search costs:

$$(1) \quad \max \int_t^\infty e^{-r(s-t)} \left(\sum_{i=1}^N \sum_{j=1}^N f_{ij}(s) m_{ij}(s) / 2 - \sum_{i=1}^N c(\rho_i(s)) u_i(s) \right) ds$$

The first term is the gross production in the economy (being careful not to double-count), while the second term gives the search costs incurred by unmatched agents.

The planner maximizes (1) by setting the search intensity function $\rho_i(t)$. In addition, the social planner faces a constraint on the evolution of the state variables:

$$(2) \quad dm_{ij}(t) \leq \left(\frac{\rho_i(t) \rho_j(t) u_i(t) u_j(t)}{\sum_{k=1}^N \rho_k(t) u_k(t)} - \delta m_{ij}(t) \right) dt \quad \text{and} \quad m_{ij}(t) \geq 0.$$

If the social planner permits all feasible (i, j) matches during any interval of time dt , inequality (2) will bind. The first term in parenthesis gives the creation rate of new matches, while the second term gives the exogenous destruction rate. But this is only a cap on the growth of $m_{ij}(t)$, since the planner may destroy any new or existing matches at will.

3 The Nonstationarity of Optimal Policies

This section considers two examples that illustrate the possibility that nonstationary policies may dominate stationary ones. The first one treats the search intensity ρ as exogenous, while the second one relies explicitly on endogenous and time-varying search intensity.

3.1 Exogenous Search Intensity

There are equal measures of two types of agents 1 and 2. They are perfectly patient, $r = 0$, and so only care about their average payoff. Type 1 agents cannot produce with other type 1 agents, $f_{11} \ll 0$, sufficiently negative that it is not socially optimal for them to match, $m_{11} \equiv 0$.⁵ ‘Good’ matches between a pair of type 2 agents are productive, with normalization $f_{22} = 2$. ‘Mixed’ matches (1, 2) produce an intermediate level of output, $f_{12} \in (0, 1)$. The restriction $f_{12} < 1$ ensures that it is better to create a good match, using two type 2 agents, than a mixed match, which only uses one high productivity agent. This implies that good matches are always allowed. The interesting question is whether mixed matches should be.

To address this, suppose for now that there is a steady state optimum, so $m_{12}(t)$ and $m_{22}(t)$ are constant. Stationarity implies $dm_{ij}(t) = 0$ in inequality (2). The permanence of good matches additionally implies the constraint is binding for $i = j = 2$. Solving this equality yields a linear relationship $m_{22} = k(\frac{1}{2} - m_{12})$, where $k = 1 - \sqrt{\delta/(\delta + \rho)}$. Additionally, evaluating inequality (2) for mixed matches yields the constraint $m_{12} \leq \frac{1}{2}k$. Together, these define the feasible steady state values of the state variables, indicated by the downward sloping line segment connecting $(0, \frac{1}{2}k)$ with $(\frac{1}{2}k, \frac{1}{2}k(1 - k))$ in Figure 1. With $f_{22} = 2$, average output is $f_{12}m_{12} + m_{22} = (f_{12} - k)m_{12} + \frac{1}{2}k$. That means that, in the class of stationary policies, it is optimal to consummate mixed matches if and only if $f_{12} > k$.

Now let us consider nonstationary paths. For expositional ease, start with the nongeneric case $f_{12} = k$, so the planner is indifferent between steady states. Moreover, suppose she opts not to allow mixed matches, driving the state to $(0, \frac{1}{2}k)$. Suddenly at time t she changes her mind, and decides to permit mixed matches. Initially, the share of mixed matches increases, $dm_{12}(t)/dt > 0$. However, there is no offsetting decrease in the share of good matches; $m_{12}(t) = 0$ implies $dm_{22}(t) = 0$ as well. The subsequent path of the economy is shown by the concave arc starting at the initial condition $(0, \frac{1}{2}k)$. Note that she never has to pay for this permanent policy change; the economy eventually converges to $(\frac{1}{2}k, \frac{1}{2}k(1 - k))$ from outside the set of (equal output) points sustainable in steady state.⁶

⁵It is not enough to assume that $f_{11} = 0$. In that case, the social planner might want to create (1, 1) matches, since that improves the matching process by getting these agents out of the way.

⁶If ever $m_{22} = k(\frac{1}{2} - m_{12})$ and $m_{12} < \frac{1}{2}k$, $dm_{22}(t) = 0$ and $dm_{12}(t) > 0$, forcing the system directly to the right. Thus the dynamics cannot cross the line segment connecting $(0, \frac{1}{2}k)$ with $(\frac{1}{2}k, \frac{1}{2}k(1 - k))$ until it reaches the end point.

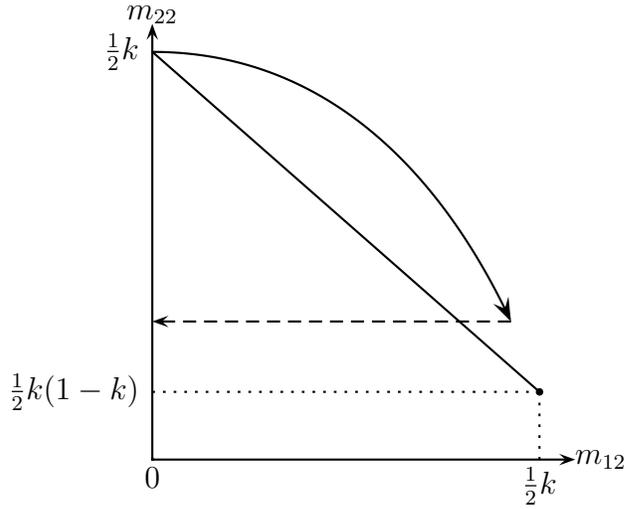


Figure 1: Nonstationary paths can move outside the steady state feasible set.

With a discount rate of zero, this permanent policy change does not raise average output, since the economy still spends almost all the time very close to the feasible steady state values. But it does clarify why nonstationary policies may be optimal. When mixed matches are prohibited at $(0, \frac{1}{2}k)$, a large number of type 1 agents clog the unmatched pool. If this makes it sufficiently hard to create good matches, it is optimal to allow mixed matches. And as the economy moves along the concave arc, the shortage of unmatched type 2 agents actually worsens, since each new mixed match removes one of each type from the searching pool. Thus mixed matches become increasing desirable over time. But suppose that after allowing these mixed matches for some time, the planner suddenly destroys them. The economy jumps back to the left, as indicated by the dashed vector. There are now a high proportion of type 2 agents in the unmatched pool, which makes it easier to create good matches, and in turn justifies the destruction of the mixed matches.

This suggests that an optimal policy should cycle between intervals in which mixed matches are permitted, and intervals during which they are prohibited. To show that such a policy can do strictly better than a stationary one, set $\rho = 4$ and $\delta = 1/20$, so $k = 8/9$.⁷ If $f_{12} = k$, the optimal steady state achieves output of $k/2 = 4/9$. But the following policy averages 3.6% higher output: For 7.56

⁷An optimal path has a limit cycle in this example. This result does not generalize to economies with more than two state variables, which may exhibit optimal chaotic behavior.

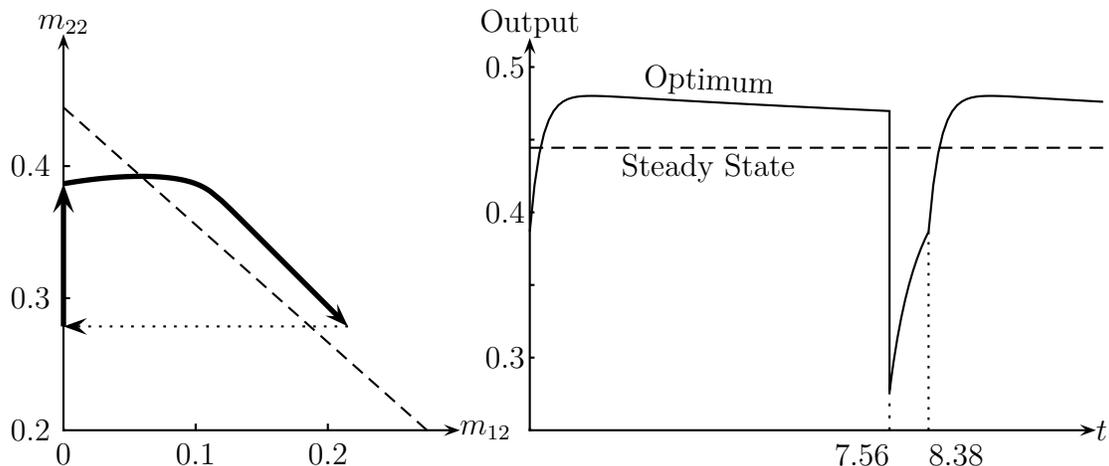


Figure 2: Optimal nonstationary policy and the consequent output dynamics.

periods, allow mixed matches; then destroy them, and for a subsequent 0.82 periods, prohibit any new mixed matches; then repeat. The state variables cycle clockwise, as indicated in the left panel of Figure 2, spending much of the time beyond the set of points achievable in steady state, as indicated by the dashed line segment. Because of this, output is often higher than in steady state, as the right panel of Figure 2 shows. Of course, this limit cycle is based on the nongeneric assumption that all steady states generate the same output, $f_{12} = k/2$. But continuity implies that similar limit cycles achieve significantly higher output than the (unique) social optimum for nearby values of f_{12} .

As the simplicity of the above example suggests, optimal nonstationary behavior seems to be an extremely general feature of search and matching models. The key reason is that matching rates that are not sustainable in steady state may be visited repeatedly on a nonstationary path. In particular, in any steady state in which some matches are rejected, there is a first order benefit to creating a few of those matches, and only a second order cost in terms of lost opportunities for other matches. This is a free lunch, and will always be optimal in a sufficiently small portion.

3.2 Endogenous Search Intensity

The type of nonstationary behavior discussed in the first example may manifest itself in a model with endogenous search intensity. But endogenous search intensity also introduces a new type of nonstationary behavior. Again assume that there are equal measures of two types of agents 1 and 2, still perfectly patient. Now

normalize $f_{11} = f_{22} = 1$, but set $f_{12} = 0$. Again first look at the class of stationary solutions to this problem. Given the symmetry of the problem, all should agents use the same search intensity ρ . In steady state, this ensures that they have the same unmatched rates u_i ; and hence that half of all meetings are with the same type of agent and thus result in matches. Then (2) determines the steady state matching rate, $m_{11} = m_{22} = \rho/2(\rho + 2\delta)$. The social planner's problem is then simply to choose ρ to maximize

$$\frac{\rho}{2(\rho + 2\delta)} - c(\rho)$$

Let ρ^* denote the maximizing value of ρ , the optimal stationary search intensity.

But nonstationary paths can do better. Fix $\varepsilon > 0$. For $t \in [0, \varepsilon) \cup [2\varepsilon, 3\varepsilon) \cup \dots$, set $\rho_1(t) = \rho^*$ and $\rho_2(t) = 0$, and at other times $\rho_2(t) = \rho^*$ and $\rho_1(t) = 0$. With this strategy, *all* meetings result in matches, since different types of agents never search at the same time. Of course, the agents only search half the time, so the *average* matching rate is essentially unchanged. Indeed, one can show that in the limit as ε converges to zero, the matched rates m_{11} and m_{22} converge to their values in the steady state discussed in the previous paragraph. But this nonstationary policy is superior: since only half the unmatched agents search at any point in time, search costs are cut in half. Time-varying policies are effectively being used to coordinate search activities.

What would an optimal policy look like?⁸ Whenever there are more unmatched type 1 agents than type 2 agents, the social planner wants type 1 agents to search; and whenever the reverse is true, she wants type 2 agents to search. Of course, when type 1 agents search, their matching rate tends to increase, while the matching rate of type 2 agents falls. Thus the optimal policy must involve ‘chattering’, i.e. the limit of the policy described above when ε converges to zero.⁹ But such a policy is not well defined, which means that there is no socially optimal policy in this model. The stationary analysis is completely misleading.

Using time-varying search intensity to coordinate search activities may seem to go against the spirit of anonymous search and matching models.¹⁰ Yet we cer-

⁸An optimal policy will always be state-, not time-dependent. It is possible to express the policy described above as an equivalent state-dependent policy, i.e. with search intensities dependent only on m_{11} and m_{22} . The policy described in this paragraph is state-dependent.

⁹Also, the optimal search intensity ρ is higher because of the reduced search costs.

tainly do not want to preclude the use of nonstationary policies in this inherently nonstationary economy. Does this represent a flaw in our specification of the search process? Coordinating search intensities is useful whenever there is complementarity between search intensities in the matching function. In our specification, borrowed from Pissarides (2000), if agent i and j each search twice as hard, the probability of a meeting between them quadruples.

Contrast this with an alternative search process proposed by Diamond (1982b) and Mortensen (1982). Unmatched agents all register with a central matching agency. An agent's search intensity determines how frequently she calls the agency. When she calls, she is given information on how to contact one other unmatched agent, selected randomly from the population registered at the agency, but independent of that agent's search intensity. Conversely, another agent may search and be given her contact information. This implies that (i, j) meetings occur at rate $(\rho_i + \rho_j)u_i u_j / \sum_{k=1}^N u_k$, linearly dependent on the search intensities. Coordinating search intensities confers no advantage in such a model world.

While this eliminates the type of nonstationary behavior described here, it relies on this being an exact description of the search process. To the extent that j 's search intensity has any positive effect on the likelihood of i contacting her, the model exhibits complementarities, and 'chattering' policies may be optimal. Thus while we feel the Diamond-Mortensen search technology may be instructive (and we discuss it further in Section 7), we are not convinced that it resolves the chattering problem. We are forced to conclude that such policies may be optimal in a reasonable, open class of search and matching models.

4 Optimum

The social planner chooses a time path of search intensities to maximize (1) subject to (2). The unusual law of motion for the control variable complicates our analysis. Rather than directly tackling the problem at hand, we consider a related problem.

¹⁰But it is consistent with observed features of the matching process, such as periodic job fairs and gay or lesbian nights in bars.

We assume m_{ij} is differentiable, with

$$\dot{m}_{ij}(t) = \frac{\rho_i(t)\rho_j(t)u_i(t)u_j(t)}{\sum_{k=1}^N \rho_k(t)u_k(t)} - x_{ij}(t)m_{ij}(t)$$

Here $x_{ij}(t) \in [\delta, \bar{x}]$ represents the rate of endogenous plus exogenous match destruction, chosen optimally by the social planner, and \bar{x} represents a (large) finite bound on match destruction. We then take the limit as \bar{x} goes to infinity, effectively allowing the planner to dissolve positive measures of matches instantaneously.

We represent this constrained problem using a current-valued Hamiltonian with multiplier $R_{ij}(t) \equiv R_{ji}(t)$ on the matched rate $m_{ij}(t) \equiv m_{ji}(t)$:

$$H = \sum_{i=1}^N \sum_{j=1}^N \left(f_{ij}m_{ij}(t)/2 + \frac{\rho_i(t)\rho_j(t)u_i(t)u_j(t)R_{ij}(t)}{\sum_{k=1}^N \rho_k(t)u_k(t)} - x_{ij}(t)m_{ij}(t)R_{ij}(t) \right) - \sum_{i=1}^N c(\rho_i(t))u_i(t)$$

There are four types of necessary first order conditions. First is the optimality condition for job destruction, which takes a bang-bang form:

$$(3) \quad x_{ij}(t) = \begin{cases} \delta \\ \bar{x} \end{cases} \text{ if } \frac{\partial H}{\partial x_{ij}(t)} = m_{ij}(t)R_{ij}(t) \geq 0$$

Note that $m_{ij}(t) \geq 0$ by definition of this state variable. In addition, when $m_{ij}(t) = 0$, the choice of x is irrelevant. Hence the optimality condition is equivalent to imposing $x_{ij}(t) = \delta$ when the match value $R_{ij}(t)$ is positive, and $x_{ij}(t) = \bar{x}$ when $R_{ij}(t)$ is negative.

The next necessary condition is the costate equation, that the partial derivative of the Hamiltonian with respect to the state variable $m_{ij}(t)$ is the flow value of the costate variable $rR_{ij}(t) - \dot{R}_{ij}(t)$. Noting that the state variable appears in the unmatched rates u_i , we obtain:

$$(4) \quad 2(r + x_{ij}(t))R_{ij}(t) - 2\dot{R}_{ij}(t) = f_{ij} + c(\rho_i(t)) + c(\rho_j(t)) - 2\rho_i(t)E_{k,t}R_{ik}(t) - 2\rho_j(t)E_{k,t}R_{jk}(t) + (\rho_i(t) + \rho_j(t))E_{k,t}E_{l,t}R_{kl}(t)$$

Here we introduce the expectations operator, which denotes the expected value of its argument drawn from the unmatched population at a particular point in time, i.e. for any type-specific variable ϕ_1, \dots, ϕ_N ,

$$E_{k,t}\phi_k \equiv \frac{\sum_{k=1}^N \rho_k(t)u_k(t)\phi_k}{\sum_{k=1}^N \rho_k(t)u_k(t)}.$$

The transversality condition is the third necessary restriction:

$$\lim_{T \rightarrow \infty} e^{-rT} R_{ij}(T) m_{ij}(T) = 0.$$

We can now integrate the costate equation (4) forward, using the transversality condition to pin down the constant of integration:

$$(5) \quad R_{ij}(t) = \int_t^\infty e^{-(r+x_{ij}(s))(s-t)} \frac{f_{ij} - v_i(s) - v_j(s)}{2} ds$$

where $v_i(t)$ is the flow value of an unmatched agent:

$$(6) \quad v_i(t) = -c(\rho_i(t)) + \rho_i(t)(2E_{k,t}R_{ik}(t) - E_{k,t}E_{l,t}R_{kl}(t))$$

To further simplify these expressions, take the limit as $\bar{x} \rightarrow \infty$. Then whenever $R_{ij}(t) < 0$, (3) implies matches are instantly terminated, allowing us to write (5) as

$$(7) \quad R_{ij}(t) = \max_{T \in [t, \infty]} \int_t^T e^{-(r+\delta)(s-t)} \frac{f_{ij} - v_i(s) - v_j(s)}{2} ds$$

The value of a match is the present discounted value of future surpluses, where discounting accounts for impatience and match impermanence, and surplus is output in excess of unmatched values. At this limit, we can also simplify the expression for the dynamics of the state variable.

$$(8) \quad \begin{aligned} R_{ij}(t) > 0 &\Rightarrow \dot{m}_{ij}(t) = \frac{\rho_i(t)\rho_j(t)u_i(t)u_j(t)}{\sum_{k=1}^N \rho_k(t)u_k(t)} - \delta m_{ij}(t) \\ R_{ij}(t) = 0 &\Rightarrow 0 \leq m_{ij}(t) \leq m_{ij}(t-) \end{aligned}$$

where $m_{ij}(t-)$ denotes the left hand limit of $m_{ij}(s)$ as $s \rightarrow t$. Note equation (7) ensures that $R_{ij}(t)$ is nonnegative.

Finally, the optimality condition for the choice of ρ is that it maximize the Hamiltonian. This is equivalent to requiring ρ to maximize the right hand side of (6), i.e.:

$$(9) \quad v_i(t) = \max_{\rho \geq 0} -c(\rho) + \rho(2E_{k,t}R_{ik}(t) - E_{k,t}E_{l,t}R_{kl}(t))$$

and $\rho_i(t)$ solves the maximization problem. Note that in the case where there is only one type of agent, (9) reduces to $v(t) = \max_{\rho} -c(\rho) + \rho R(t)$. The flow value of an unmatched agent is the probability that she meets someone and has the opportunity to create social value R , minus the search costs. More generally, the flow value consists of three terms. First, an unmatched agent i is costly, because she searches. Second, by searching, she may meet a type k agent, drawn randomly from the population, and create a new match with value R_{ik} . The ‘2’ in this term represents the thick market externality, since she helps out her new partner k as well as herself. It is increasing in the expected value created by agent i . Finally, at the same time, she makes it more difficult for type k and l agents to meet, which imposes a social cost, depending on *their* values while matched. This last term represents the congestion externality. It is linear in agent i ’s search intensity, but otherwise independent of that agent’s activity.

For a given path of the state variables, (7) and (9) determine the value of matches and of unmatched agents. This allows us to provide a concise characterization of a social optimum:

Proposition 1. *If a tuple (v, R, ρ, m) is a social optimum, then: R solves (7) given v ; v solves (9) given R, ρ , and m ; $\rho_i(t)$ solves the right hand side of (9) given R, ρ , and m ; and m solves the state equations (8) given some initial conditions, R , and ρ .*

5 Equilibrium

In a decentralized search equilibrium, an individual agent seeks to maximize the expected present value of her income net of search costs. Her strategy includes her search intensity and matching decisions. Following the search literature, we assume matched agents divide their output f_{ij} according to the Nash bargaining solution, defined precisely below. This implies that all bilateral gains from trade are exploited, and in particular that a pair breaks up if and only if it is in the

interest of both agents. Nevertheless, agents ignore the effect of their actions on third parties, e.g. potential partners, which drives a wedge between equilibrium and optimal behavior. To facilitate comparing the decentralized equilibrium and social optimum, we introduce a search tax into this environment: each unmatched agent pays a flow tax in proportion to her search intensity, $\tau_i(t)\rho_i(t)$.

With this additional notation, we characterize an equilibrium using recursive equations. Start with the expected value of an agent i matched with an agent j at time t , $W_{i|j}(t)$, expressed as a function of future unmatched values $W_i(s)$, $s \geq t$:

$$(10) \quad W_{i|j}(t) = \max_{T \in [t, \infty]} \int_t^T e^{-(r+\delta)(s-t)} (\pi_{i|j}(s) + \delta W_i(s)) ds + e^{-(r+\delta)(T-t)} W_i(T)$$

An (i, j) match ends at an optimally chosen future date $T \in [t, \infty]$, when it is in the mutual interest of i and j . Until then, i gets an endogenous flow payoff $\pi_{i|j}$, and suffers exogenous match destruction with flow probability δ , which leaves her unmatched. If the match survives until date T , she gets her unmatched value $W_i(T)$, discounted back to date t .

We can express (10) in a more useful form by noting that for any $T \in [t, \infty]$,

$$W_i(t) \equiv \int_t^T e^{-(r+\delta)(s-t)} ((r + \delta)W_i(s) - \dot{W}_i(s)) ds + e^{-(r+\delta)(T-t)} W_i(T),$$

as can be verified using integration by parts. Subtracting this from (10) implies

$$(11) \quad S_{ij}(t) \equiv W_{i|j}(t) - W_i(t) = \max_{T \in [t, \infty]} \int_t^T e^{-(r+\delta)(s-t)} (\pi_{i|j}(s) - w_i(s)) ds$$

where $w_i(s) \equiv rW_i(s) - \dot{W}_i(s)$ is the flow value of an unmatched agent. $S_{ij}(t)$ is the surplus that i gets from being matched with j at time t .

Since a pair always has the option to destroy the match immediately, $S_{ij}(t)$ is nonnegative. If it is positive, there are gains from trade, which are divided up according to a dynamic Nash bargaining solution.¹¹ The threat of each agent is to destroy the match, making $W_i(t)$ and $W_j(t)$ the relevant threat points, while the payoff to be divided up is $W_{i|j}(t) + W_{j|i}(t)$. Thus Nash bargaining imposes $S_{ij}(t) = S_{ji}(t)$. Equivalently, the integrand of (11) must be equal for each party to

¹¹Note that agents bargain over the expected present value of payoffs, not current payoffs. Coles and Wright (1998) is the first application of this bargaining rule in a search model.

the match,

$$\pi_{i|j}(s) - w_i(s) = \pi_{j|i}(s) - w_j(s) = f_{ij} - \pi_{i|j}(s) - w_j(s)$$

where the second equality uses the fact that output f_{ij} is divided between i and j . This pins down individual flow payoffs π . Substituting into (10) implies

$$(12) \quad S_{ij}(t) = \max_{T \in [t, \infty]} \int_t^T e^{-(r+\delta)(s-t)} \frac{f_{ij} - w_i(s) - w_j(s)}{2} ds$$

Each agent's share of surplus in an (i, j) match is half the the present discounted value of future flow surplus, which is output in excess of the sum of the flow unmatched values. Since i and j match when this surplus is positive, the state variable follows a standard law of motion, analogous to condition (8):

$$(13) \quad \begin{aligned} S_{ij}(t) > 0 &\Rightarrow \dot{m}_{ij}(t) = \frac{\rho_i(t)\rho_j(t)u_i(t)u_j(t)}{\sum_{k=1}^N \rho_k(t)u_k(t)} - \delta m_{ij}(t) \\ S_{ij}(t) = 0 &\Rightarrow 0 \leq m_{ij}(t) \leq m_{ij}(t-) \end{aligned}$$

Finally, we close the system by writing the recursive equation for the flow value of an unmatched agent:

$$(14) \quad rW_i(t) - \dot{W}_i(t) \equiv w_i(t) = \max_{\rho \geq 0} -c(\rho) - \tau_i(t)\rho + \rho E_{j,t} S_{ij}(t)$$

An unmatched agent must choose her search effort. Her flow payoff comes from the search cost and search tax, $-c(\rho) - \tau_i\rho$, while the probability of meeting a new potential partner is ρ , and the expected capital gain is $E_{j,t} S_{ij}(t)$.

A precise definition of a search equilibrium follows from these expression:

Definition 1. *A search equilibrium is any tuple (w, S, ρ, m) such that: S solves (12) given w ; w solves (14) given S , ρ , and m ; $\rho_i(t)$ solves the right hand side of (14) given S , ρ , and m ; and m solves the state equations (13) given some initial conditions, ρ , and S*

6 Search Externalities

We are now in a position to characterize the search externalities in this economy. Our characterization relies on the following Pigouvian tax:

Proposition 2. *Let $\{v, R, \rho, m\}$ be a social optimum. Impose a search tax*

$$(15) \quad \tau_i^*(t) \equiv E_{k,t} E_{l,t} R_{kl}(t) - E_{k,t} R_{ik}(t)$$

where expectations are taken with respect to the search intensities and unmatched rates that prevail at the social optimum. Then there is a decentralized equilibrium that is socially optimal.

Proof. The optimality of the tax can immediately be confirmed by substituting for τ in (14) and comparing the characterization of a social optimum in Proposition 1 with the description of a search equilibrium in Definition 1. \square

We do not claim that this search tax can be implemented in practice. Instead, it informs us as to the nature of search externalities.

Total tax revenue is always equal to zero, $E_{i,t} \tau_i^*(t) = 0$. This implies that with only one type of agent, $N = 1$, the decentralized equilibrium is efficient in the absence of any taxes. This conclusion is in accordance with Hosios (1990). Although each agent only captures half the surplus from a match, which might seem to be the hallmark of underinvestment, this disincentive to search internalizes the fact that if one agent searches harder, it becomes more difficult for other agents to find partners. The congestion and thick-market externalities cancel.

In contrast, as long as there is real heterogeneity in the model, a search equilibrium is not socially optimal in the absence of taxes. This is easiest to see in the case where the social optimum has a steady state.

Proposition 3. *Suppose the social optimum is stationary. Assume the production function is weakly monotonic, $f_{ik} \leq f_{jk}$ whenever $i < j$. Then higher types pay weakly lower Pigouvian taxes, $\tau_i^* \leq \tau_j^*$ whenever $i < j$.*

If in addition the search cost function is nonnegative valued, $f_{Nk} > f_{1k}$ for all k , and $f_{NN} > 0$, then the highest type gets a search subsidy and the lowest type pays a search tax: $\tau_N^ < 0 < \tau_1^*$.*

Note that in many examples, the social optimum asymptotes towards a steady state. Thus it is not vacuous to discuss a stationary social optimum.

Proof. A stationary social optimum can be characterized by stationary versions of (7) and (9):

$$R_{ik} = \max \left\langle \frac{f_{ik} - v_i - v_k}{2(r + \delta)}, 0 \right\rangle$$

$$v_i = \max_{\rho \geq 0} -c(\rho) + \rho(2E_k R_{ik} - E_k E_l R_{kl})$$

Fix $i < j$, and suppose in order to find a contradiction that $E_k R_{ik} > E_k R_{jk}$. Then the second equation implies $v_i \geq v_j$. The first equation, with $f_{ik} \leq f_{jk}$ implies $R_{ik} \leq R_{jk}$ for all k . Taking expectations yields the desired contradiction, proving that $E_k R_{ik} \leq E_k R_{jk}$. Weak monotonicity of the Pigouvian tax then follows directly from its definition (15).

Now suppose $v_1 = v_N$. With $f_{1k} < f_{Nk}$ for all k , the only way that $R_{1k} = R_{Nk}$ is if both are equal to zero. This is true in expectations as well: either $E_k R_{1k} < E_k R_{Nk}$ or $E_k R_{1k} = E_k R_{Nk} = 0$. The latter possibility corresponds to the case when there is no value in any match, which is ruled out by the technical assumptions: if $E_k R_{Nk} = 0$, nonnegativity of the search cost function c implies $v_N \leq 0$; but then since $f_{NN} > 0$, $R_{NN} > 0$ as well, contradicting $E_k R_{Nk} = 0$. We conclude that the expected surplus function satisfies $E_k R_{1k} < E_k R_{Nk}$, and so (15) implies $\tau_1^* > \tau_N^*$. Finally, since $E_k \tau_k^* = 0$, zero must lie in between these extremal values. \square

The nondegenerate search taxes shed light on the nature of externalities. Take an agent i who should receive a subsidy, $\tau_i^* < 0$. This would have two effects on her behavior. Most obviously, she would search harder than she does without the subsidy. But in addition, the subsidy would reduce the cost of being unmatched. To see this, suppose that she did not change her search intensity in response to the subsidy. Then it would simply raise the value of being unmatched, eliminating the surplus in previously marginal matches. Conversely, a search tax $\tau_i^* > 0$ would reduce search intensity and make agents more willing to match. Note that the implication for equilibrium unmatched rates u_i is ambiguous, since the search and matching inefficiencies work in opposite directions.

Put differently, in a decentralized steady state equilibrium without search subsidies, the most productive agents do not search hard enough and are too willing to

accept a match when they meet another agent. This is the effect of the thick-market externality: when a productive agent searches harder or rejects a match, she confers a benefit on other agents who have an opportunity to meet her. A search subsidy internalizes this effect. The opposite inefficiency characterizes the least productive agents. They impose a congestion externality on others: when an unproductive agent searches harder or rejects a match, the benefits of search are diminished for other agents, who must waste some of their time meeting her. A search tax forces her to internalize this effect.

Proposition 3 does not easily generalize to handle nonstationary social optima. Consider the following example: $f_{11} = 100$, $f_{12} = 101$, and $f_{22} = 108$, with the usual exogenous search frictions, $\delta = 1/20$, $\rho = 4$, and $r = 0$. Also assume that most of the agents have low productivity, $\ell_1 = 2/3$. Then in an optimal policy, agents always match with their own type, but cycle through 3.51 periods in which mixed matches are acceptable, followed by 0.91 periods in which they are not. Because there are more low productivity agents, unmatched high productivity agents are scarce, particularly towards the end of the periods in which mixed matches are acceptable. That means most new meetings are with low productivity agents. For low productivity agents, that is not too bad, since low productivity matches at least have the benefit of lasting forever. But for high productivity agents, searching is very unproductive at these times. Most of their meetings result in mixed matches, and the surplus in mixed matches R_{12} is very small, since they will be destroyed in the near future. This pulls down the flow value of high productivity agents, $v_2(t)$, and for the last twenty percent of the mixed-matching phase, it is lower than the flow value of low productivity agents $v_1(t)$. That implies that over the same interval, high productivity agents must optimally pay a search tax and low productivity agents receive a search subsidy.

Still, there is a sense in which higher types receive larger *average* search subsidies, even in nonstationary environments, at least when search intensity is exogenous and common across agents. Monotonicity of the production function f implies monotonicity of the social *total* value function $V_i(t) = \int_t^\infty e^{-r(s-t)} v_i(s) ds$, a result that follows from the ability of high types to mimic the behavior of low types. Thus

for $i < j$,

$$\begin{aligned} 0 \leq \frac{V_j(t) - V_i(t)}{2\rho} &= \int_t^\infty e^{-r(s-t)} E_{k,t}(R_{jk}(s) - R_{ik}(s)) ds \\ &= - \int_t^\infty e^{-r(s-t)} (\tau_j(s) - \tau_i(s)) ds \end{aligned}$$

where the final equality follows from the optimal tax (15). That is, the average value of future search taxes, discounted at the market interest rate, is always lower for higher types.

7 Discussion

An important question is how dependent our results are on our specification of the search process. As we discussed in Section 3.2, the literature offers at least one alternative linear search technology (Diamond 1982b, Mortensen 1982).¹² An agent who searches with intensity ρ contacts a type j agent at rate $\rho u_j / \sum_{k=1}^N u_k$, and regardless of her own search activity, a type j agent contacts her at rate $\rho_j u_j / \sum_{k=1}^N u_k$. With homogeneous agents, equilibrium search behavior is optimal if an agent keeps all the surplus from the matches that she initiates (Mortensen 1982). An agent's search does not affect the ability of other agents to contact her, and so she should not receive any of the benefits from those meetings. However, she incurs the full cost of initiating contact with other agents, and so must get the full benefit if she is to search efficiently.

This finding does not carry over to an environment with heterogeneous agents, however.¹³ When an agent decides to match, she alters the distribution of un-

¹²Diamond (1982a) considers a related search process. Unmatched agents contact others at a rate equal to their search effort; however, the other agent is drawn randomly from the entire population, matched and unmatched. If the agent is already matched, matching is impossible. Thus an agent who searches at rate ρ contacts a type j agent at rate ρu_j , and is contacted by such an agent at rate $\rho_j u_j$. This is an example of a quadratic search technology, since (assuming ρ is constant and equal across agents) the flow of matches depends on the square of the aggregate unmatched rate $\sum_{k=1}^n u_k$.

With this search process, the decision of one agent to remain unmatched does not make it more difficult for other agents to find each other; there is no congestion externality. However, there is still a thick-market externality. Regardless of how surplus is divided, an agent fails to account for the fact that when she matches, future meetings are given up. Diamond shows that in a homogeneous agent environment, efficiency demands a search subsidy. This result carries over to a model with heterogeneous agents.

matched agents. This changes the returns to search for other agents, an external effect that is not internalized by Mortensen's (1982) proposed allocation of property rights. In particular, when a high productivity agent matches, the unmatched pool becomes less attractive. Thus from a social perspective, these agents are generally too willing to match, and a search subsidy is desirable. Conversely, low productivity agents are too reluctant to match. This is consistent with the findings from our model.

We conclude that a necessary condition for efficiency of a search equilibrium with heterogeneous agents, is that one agent's behavior must not alter the opportunities available to other agents. Since a fundamental characteristic of these models is that matching decisions affect search opportunities by altering the distribution of unmatched agents, this is a stringent requirement. In general, high productivity agents are too willing to match, because they ignore the benefit they impart on the unmatched distribution. Similarly, low productivity agents fail to internalize the congestion that they impose on search markets. These external effects can only be appreciated in a model with heterogeneous agents and a nontrivial matching decision.

References

- Acemoglu, Daron and Robert Shimer**, "Holdups and Efficiency with Search Frictions," *International Economic Review*, 1999, 40 (4), 827–849.
- Burdett, Kenneth and Melvyn Coles**, "Separation Cycles," *Journal of Economic Dynamics and Control*, 1998, 22 (7), 1069–1090.
- Coles, Melvyn and Randall Wright**, "A Dynamic Equilibrium Model of Search, Bargaining, and Money," *Journal of Economic Theory*, 1998, 78 (1), 32–54.
- Diamond, Peter**, "Aggregate Demand Management in Search Equilibrium," *Journal of Political Economy*, 1982, 90 (5), 881–894.
- , "Wage Determination and Efficiency in Search Equilibrium," *Review of Economic Studies*, 1982, 49 (2), 217–227.

¹³The formal proof of these results has a similar structure to the analysis in the paper. Details are available upon request.

- Hosios, Arthur**, “On the Efficiency of Matching and Related Models of Search and Unemployment,” *Review of Economic Studies*, 1990, 57 (2), 279–298.
- Mortensen, Dale**, “Property Rights and Efficiency in Mating, Racing, and Related Games,” *American Economic Review*, 1982, 72 (5), 968–979.
- Pissarides, Christopher**, *Equilibrium Unemployment Theory*, second ed., Cambridge, MA: MIT Press, 2000.
- Sattinger, Michael**, “Search and the Efficient Assignment of Workers to Jobs,” *International Economic Review*, 1995, 36 (2), 283–302.
- Shi, Shouyong**, “Frictional Assignment,” 1998. Mimeo.
- Shimer, Robert and Lones Smith**, “Assortative Matching and Search,” *Econometrica*, 2000, 68 (2), 343–370.