Directed Search and Competitive Search Equilibrium: A Guided Tour

Randall Wright, Philipp Kircher, Benoît Julien, and Veronica Guerrieri

This essay surveys the literature on directed search and competitive search equilibrium, covering theory and a variety of applications. These models share features with traditional search theory, but also differ in important ways. They share features with general equilibrium theory, but with explicit frictions. Equilibria are often efficient, mainly because markets price goods plus the time required to get them. The approach is tractable and arguably realistic. Results are presented for finite and continuum economies. Private information and sorting with heterogeneity are analyzed. While emphasizing issues and applications, we also provide several hard-to-find technical results. (JEL D50, D83)

1. Introduction

Search theory contributes significantly to fundamental and applied research, and is relevant for understanding many phenomena that are troublesome for classical economics. Examples include the coexistence of unemployment and vacancies; price or wage dispersion and stickiness; bid-ask spreads; the difficulties of bilateral trade that generate a role for money and intermediation; partnership formation; long and variable durations in the time to execute trades in labor, housing, and other markets; et cetera. This essay surveys a branch of the area concentrating on directed search and competitive search...
equilibrium. While the plan in what follows is to go into considerable detail on the literature, to give a hint up front, we mention influential papers by Peters (1984, 1991), Montgomery (1991), Shimer (1996), and Moen (1997).1

To be precise, for our purposes, an economic model has two components: an environment, including descriptions of the set of agents with preferences and technologies; and a mechanism or solution concept mapping environments onto outcomes. Directed search is a feature of the environment. To explain this, first note that in search theory agents are modeled as trading with each other, and often bilaterally, different from Walrasian theory, where they simply trade with (slide along) their budget lines.2 Directed search means agents see some, although perhaps not all, characteristics of other agents, and based on that choose where to look for counterparties. This contrasts with random search, where meetings are exogenous. The characteristics of agents in most of the models described below include their posted terms of trade—generally contracts, although sometimes these are simply prices. This contrasts with traditional search models that assume agents bargain after they meet, and also with those that assume price posting when the posted terms do not influence who meets whom.

While directed and random search are different approaches that can be applied in single-agent decision theory, competitive search equilibrium is solution concept mapping environments into outcomes. While there is not complete consensus on usage, we take a stand: competitive search equilibrium means that agents on one side of the market post the terms of trade, while agents on the other side observe what is posted and direct their search accordingly. This can be applied in environments with finite or infinite numbers of agents; the situation where these numbers are sufficiently large that certain strategic considerations can be ignored is called perfectly competitive search equilibrium.

Consider any two-sided market with, for example, buyers and sellers, firms and workers, or borrowers and lenders, trying to get together in pairs. Traditional search assumes they meet exogenously, at random, although whether a meeting results in trade can be endogenous. Directed search is different because agents use the information to target their search toward particular types, and sometimes even particular individuals. This can be based on primitive characteristics of agents—for instance, buyers can search for sellers of particular goods—and on endogenous characteristics like the posted terms of trade. In competitive search, equilibrium resources are allocated through the terms of trade plus the probability of trade. This is distinct from Walrasian theory, where trading probabilities play no role, and from traditional random search models, where prices have a relatively minor allocative role.3

Competitive search thus integrates elements of general equilibrium (GE) theory and

---

1 Highlighting a few papers like this was suggested by a referee, but it is not easy to know where to draw the line. Other early related work includes Sattinger (1990) and Hosios (1990); one might also say the papers reviewed in section 5 were early and foundational; etc.

2 Some search models do have Walrasian pricing, for example, Lucas and Prescott (1974) and Rocheteau and Wright (2005) in labor and goods markets, respectively, motivated by saying agents meet in large groups and not bilaterally.

3 As Hosios (1990) says about labor,

Though wages in bargaining models are completely flexible, these wages have nonetheless been demurred of any allocative or signaling function: this is because matching takes place before bargaining and so search effectively precedes wage-setting…. In conventional market situations, by contrast, firms design their wage offers in competition with other firms to profitably attract employees; that is, wage setting occurs prior to search.
traditional search theory, yet is still tractable and often delivers clean results.

Further, in terms of realism, Howitt (2005) puts it like this:

In contrast to what happens in [random] search models, exchanges in actual market economies are organized by specialist traders, who mitigate search costs by providing facilities that are easy to locate. Thus when people wish to buy shoes they go to a shoe store; when hungry they go to a grocer; when desiring to sell their labor services they go to firms known to offer employment. Few people would think of planning their economic lives on the basis of random encounters.

Even more colorfully, Hahn (1987) says “someone wishing to exchange his house goes to estate agents or advertises—he does not, like some crazed particle, wait to bump into a buyer.” And Prescott (2005) says “I think the bilateral monopoly problem has been solved. There are stores that compete. I know where the drug store and the supermarket are, and I take their posted prices as given. If some supermarket offers the same quality of services and charges lower prices, I shop at that lower price supermarket.”

Whether or not they realize it, while attempting to critique traditional search theory, these commentators with their pithy remarks are all describing facets of directed or competitive search. Our view is that there is a role for random matching models, with bargaining or other means for determining the terms of trade, but it is also good to know the alternatives. The models presented below provide a class of alternative that is becoming increasingly popular and has proved useful in many applications. After discussing details of the various models and applications, we will say more about the advantages and disadvantages of the alternative approaches.

We present models of finite markets and limiting results for large markets. In all of these models frictions take center stage, even when the set of agents is large. In particular, some sellers can have few customers, while others have more than they need, leading to rationing, unsold inventories, the coexistence of vacancies and unemployment, etc. The theory captures a powerful idea: if you post more favorable terms, customers may come to you with a higher probability, but not necessarily with probability 1. If a restaurant only has a certain number of tables, or a firm only wants to hire a certain number of workers, it may not be smart to go where everyone else is going. These kinds of capacity constraints play a major role in the models surveyed below, and mean that agents must consider both the terms of trade and the probability of trade. Of course capacity also matters in GE theory, but only at the market level—one cannot ask about the capacity of an agent’s trading partner, because the agents trade with the market and not with each other.

The paper is organized as follows. Section 2 starts with static models, one framed in terms of goods and one in terms of labor markets. Section 3 embeds these in dynamic GE to discuss the time to execute trades (e.g., how long an unemployed worker or unsold house remains on the market) as well as endogenous price dispersion and stickiness. Section 4 presents applications in monetary economics, where the framework provides a very natural approach. While sections 2–4 use large numbers of agents, section 5 analyzes finite markets and then takes limits as the market gets large. Section 6 studies heterogeneity and sorting. Sections 7 and 8 consider private information and mechanism design. Sections 9 and 10 mention other topics and conclude. In general, applications that do not pique a reader’s interest can be skipped without loss of continuity. As a rule of thumb, footnotes contain optional material (e.g., additional citations or technical details) and can
also be skipped without loss of continuity. Appendices, as usual, can be skipped, too, but they contain some new or hard-to-find technical material that is potentially useful.

2. Benchmark Models

2.1 Goods Markets

Consider a market with large numbers of two types of agents, called buyers and sellers, with measures \( N_b \) and \( N_s \), and let \( N = N_b/N_s \) denote the population buyer/seller ratio. One can think of buyers as households or consumers and sellers as retailers, but other interpretations are possible (e.g., producers buying inputs from suppliers). There are two tradable objects. There is an indivisible good \( q \), and sellers can produce exactly one unit at cost \( c \geq 0 \), while buyers want to consume exactly one unit for utility \( u > c \); and there is a divisible good \( x \) that anyone can produce at cost \( C(x) = x \) and consume for utility \( U(x) = x \). The idea is that there are gains from trade in \( q \), while \( x \) serves as a payment instrument that buyers use to compensate sellers. This can be interpreted as `direct barter', although more typically in the literature it is called transferable utility.

For now each seller simply posts a price \( p \), the amount of \( x \) buyers must pay to get \( q \). Each buyer directs his search after observing all posted prices (all we actually need is that each buyer observes at least two prices; see Acemoglu and Autor 2016, theorem 13.4). Suppose traders meet pairwise. Thus, if a set of buyers with measure \( n_b \) direct their search toward a set of sellers with measure \( n_s \), the probability a seller meets a buyer is \( \alpha_s = \alpha(n) \), where \( n = n_b/n_s \) is the buyer/seller ratio, also called the queue length or market tightness. Similarly, the probability a buyer meets a seller is \( \alpha_b = \alpha(n)/n \). It is standard to assume \( \alpha_s = \alpha(n) \) is increasing and concave, which implies \( \alpha_b = \alpha(n)/n \) is decreasing, given the natural restriction \( \alpha(0) = 0 \). In static or discrete-time models \( \alpha_b \) and \( \alpha_s \) are probabilities, so we impose \( 0 \leq \alpha_j \leq 1 \); in continuous-time they are arrival rates, so we only impose \( \alpha_j \geq 0 \). We also usually assume differentiability, and sometimes \( \lim_{n \to 0} \alpha'(n) = \infty \).

To understand this, imagine any two-sided market with \( n_1 \) and \( n_2 \) agents on each side, where the number of bilateral meetings between types 1 and 2 is \( \mu = \mu(n_1,n_2) \). Analogous to a production function mapping inputs into output, \( \mu \) is assumed to be increasing and concave, and usually to display constant returns to scale (CRS). Then \( \alpha_1 = \mu(n_1,n_2)/n_1 = \mu(n,1)/n \), where \( n = n_1/n_2 \), and \( \alpha_2 = \mu(n_1,n_2)/n_2 = \mu(n,1) \). This generalizes models of one-sided markets (e.g., Diamond 1982), and is more interesting because the \( \alpha \) terms depend on tightness, even with CRS. In addition, with a two-sided specification it is natural to endogenize tightness by allowing entry by one side. For now, a buyer seeks a seller with a particular \( p \), but whether he finds one is random (in section 5 an agent finds a counterparty for sure, but may or may not trade, due to capacity constraints).

A set of sellers posting the same \( p \) and buyers searching for them constitutes a submarket with tightness \( n = n_b/n_s \). Thus, a submarket is characterized by \((p,n)\). Buyers’ and sellers’ payoffs are denoted \( V_b \).
and \( V_s \). Sellers maximize \( V_s \) by posting \( (p, n) \), although it is not crucial that they post \( n \)—sellers can equivalently post only \( p \) and let buyers work out the equilibrium \( n \) for themselves. In any case, for a seller to be in business, \( (p, n) \) must deliver to buyers their market payoff \( V_b \), which is an equilibrium object, but taken as given by individuals. This is called the market utility approach.\(^6\)

Then sellers solve the following problem:

\[
(1) \quad V_s = \max_{p, n} \alpha(n)(p - c)
\]

subject to

\[
\frac{\alpha(n)}{n}(u - p) = V_b.
\]

Sellers’ payoff in a submarket is their trading probability times the surplus \( S_s = p - c \), and buyers’ payoff is their trading probability times the surplus \( S_b = u - p \). While a more rigorous definition of competitive search equilibrium appears in section 6, the idea is basically optimization and market clearing: sellers maximize \( V_s \) subject to buyers getting \( V_b \); and the \( n \) emerging from (1) is consistent with the set of buyers and sellers in the market. Notice sellers can get the same \( V_s \) from lower \( p \) if \( n \) is higher, and buyers can get the same \( V_b \) from higher \( p \) if \( n \) is lower. These trade-offs are a quintessential element of the theory.

To solve (1), use the constraint to eliminate \( p \) in the objective function:

\[
(2) \quad V_s = \max_n \{ \alpha(n)(u - c) - nV_b \}.
\]

This problem has a unique solution.\(^7\) If it is interior it satisfies the FOC \( \alpha'(n)(u - c) = V_b \). Then, given \( n \), the constraint yields \( p \) uniquely, so any active submarkets must have the same \( (p, n) \). Hence, given CRS, we only need one submarket.

There are two standard ways to proceed. The first is to assume \( N_b \) and \( N_s \) are fixed. Then the equilibrium buyer–seller ratio must be the same as the population ratio, \( n = N \) (market clearing). The FOC then implies \( V_b = \alpha'(N)(u - c) \), and the constraint implies \( p = u - NV_b/\alpha(N) \), or

\[
(3) \quad p = \varepsilon c + (1 - \varepsilon)u,
\]

where \( \varepsilon = \varepsilon(n) \equiv n\alpha'(n)/\alpha(n) \) is the elasticity of \( \alpha(n) \) with respect to tightness. Hence, price is a weighted average of cost and utility that splits the ex post (after meeting) surplus \( S = u - c \) according to \( S_b = \varepsilon(u - c) \) and \( S_s = (1 - \varepsilon)(u - c) \). The ex ante (before meeting) payoffs can now be written \( V_b = \alpha_b\varepsilon(u - c) \) and \( V_s = \alpha_s(1 - \varepsilon)(u - c) \). This uniquely pins down the equilibrium \( (p, n, V_b, V_s) \).

The second way to proceed is to assume one side has a cost to participate, and therefore, in general, only some of them enter the market. Suppose it is sellers that have a participation cost \( k_s \). Then in equilibrium, as long as \( k_s \) is neither too big nor too small relative to \( N_s \), some but not all sellers enter, and we have the free entry condition \( V_s = k_s \). As above, the FOC implies \( V_b = \alpha'(n)(u - c) \) and the constraint implies \( p = \varepsilon c + (1 - \varepsilon)u \).

Now \( k_s = V_s = \alpha(n)(p - c) \), from which

\[\text{7Appendix A proves this in a generalized version without transferable utility: if a buyer makes payment } p \text{ to a seller, the latter gets } \nu(p) \text{ while former gets } -\gamma(p); \text{ in the text here } \nu(p) = \gamma(p) = p. \text{ In fact, what is shown is that the second-order conditions (SOCs) hold at any solution to the first-order conditions (FOCs); so if there is an interior solution it is unique, but one should also check for corner solutions.}\]
we get \( n \). Once again these conditions uniquely pin down \( (p, n, V_b, V_s) \).

**Figure 1**, a version of which appears in Peters (1991), shows the “Edgeworth box” in \((p, n)\) space. Indifference curves for buyers slope down, because they are willing to pay higher \( p \) if \( n \) is lower, so they can trade faster. Similarly, sellers are willing to accept lower \( p \) if \( n \) is higher. As in elementary microeconomics, efficient outcomes are points of tangency, tracing out the contract curve \( C \).

The left panel depicts the case without entry, where \( C \) crosses \( n = N \); the right depicts the case with entry by sellers, where \( C \) crosses the indifference curve \( V_s = k \).

For comparison, consider this problem:

\[
(4) \quad V_b = \max_{p,n} \frac{\alpha(n)}{n} (u - p)
\]

subject to

\[
\alpha(n)(p - c) = V_s,
\]

where it looks like buyers post and sellers search. It yields the same \( (p, n, V_b, V_s) \) as \( (1) \), with \( n \) fixed or with entry. One can also consider a third version, with third parties, called market makers, designing submarkets by posting \( (p, n) \) to attract buyers and sellers (see Moen 1997 and Mortensen and Wright 2002 for more discussion). Again the outcome is the same. Hence, it does not matter if buyers, sellers, or market makers post here (this is not always true; see below).

**Figure 2** describes competitive search equilibrium by depicting the solution to \( (4) \) as a “demand” for sellers \( N_s \) as a function of the “cost” \( V_s \) (given \( N_b \), choosing \( n \) is the same as \( N_s \)). One can check “demand” is decreasing and, as shown, hits 0 at finite \( V_s \). Without entry, in the left panel “supply” is vertical and equilibrium determines \( V_s \). With entry by sellers, in the right panel “supply” is horizontal at \( k_s \) and equilibrium determines \( N_s \). Indeed, one could nest these with a general upward-sloping “supply” curve by letting \( k_s \) vary with the number of homogeneous entrants, or across heterogeneous potential entrants. The point

---

**Source:** Based on Peters (1991).
of figure 2, like figure 1, is that the theory can be described using tools from elementary microeconomics.

Now consider a planner’s problem with endogenous participation by sellers,

\[
\max_n \left\{ \frac{\alpha(n)(u - c)}{n} - k_s \right\}.
\]

The first term is the expected surplus per buyer; the second is the entry cost of sellers per buyer, since \( n_s/n_b = 1/n \). If we eliminate \( p \) from the objective function in (4) using the constraint, this is the same as (5) with \( V_s = k_s \). Equilibrium is efficient.

For yet another comparison, consider bargaining instead of posting. Thus, after a buyer and seller meet, they determine \( p \) by generalized Nash bargaining,

\[
\max_{\theta} (u - p)^\theta (p - c)^{1-\theta},
\]

where \( \theta \) is buyer bargaining power. The solution is \( p = \theta c + (1 - \theta)u \), which is the same as \( p \) under posting, and hence efficient, if and only if \( \theta = \varepsilon \). This is the well-known Hosios (1990) condition: efficiency obtains if and only if agents’ bargaining powers are equal to the elasticity of the meeting technology with respect to their participation. Hence, sellers should get a share of \( u - c \) commensurate with their contribution to matching.

\[8\] The easiest interpretation is that there is no communication outside of meetings, so agents cannot post terms to attract counterparties. Another is that agents cannot commit to what they post, although that need not mean it is irrelevant (Menzio 2007; Doyle and Wong 2013; Dutu 2013; Kim and Kircher 2015; Stacey 2016, 2019).

\[9\] Earlier discussions of efficiency in related models include Mortensen (1982a, b) and Pissarides, Layard, and Hellwig (1986). More recently, Mangin and Julien (2021) show this: with one-side heterogeneity, on the side searching, several environments generate an expected trade surplus that endogenously depends on tightness; they derive a generalized Hosios condition that implements (usually) efficiency by trading off the probability of trade with not only the terms of trade, but also the expected surplus.
Since this is exactly what competitive search delivers, it is often said that it induces the Hosios condition endogenously.

If \( n = N \) is fixed, one can check \( \partial V_s / \partial N > 0 \) and \( \partial V_b / \partial N < 0 \), while \( \partial p / \partial N \approx -\varepsilon' \) where \( “a \approx b” \) means “\( a \) and \( b \) have the same sign.” Now \( \varepsilon' < 0 \), and hence \( \partial p / \partial N > 0 \), for many common meeting technologies, but not all.\(^{10}\)

Does \( \partial p / \partial N < 0 \) make sense? Yes. First note that higher \( N \) always increases \( V_s \) and decreases \( V_b \), where these can change due to either changes in \( p \) or in the trading probabilities. By construction \( \alpha(n) \) goes up and \( \alpha(N)/N \) down with \( N \), but if they move a lot, \( p \) must go down so the changes in \( V_s \) and \( V_b \) are not too big. Hence an increase in demand along the extensive margin (higher \( n \)) can lower \( p \), although one can show an increase along the intensive margin (higher \( u \)) implies \( \partial p / \partial u > 0 \) unambiguously. Similarly, with seller entry, higher \( k_b \) reduces \( N_b \) and raises \( n \), also implying \( \partial p / \partial k_b \approx -\varepsilon' \) and \( \partial p / \partial u > 0 \). That \( p \) might fall when the buyer–seller ratio rises reflects the idea that resource allocation is guided by both prices and probabilities.

2.2 Labor Markets

Now let households be sellers of their time, and firms buyers. Each firm wants to hire exactly one worker, while each household wants to land one job. Thus, \( n \) is the vacancy–unemployment ratio. Again, it does not matter here who posts and who searches. Consider a version of (1) that maximizes workers’ payoffs,

\[
(7) \quad V_s = \max_{u,n} \alpha(n)(w-b)
\]

subject to

\[
\frac{\alpha(n)}{n}(y-w) = V_b,
\]

where \( y \) is output per worker and \( b \) is the value of unemployment benefits, leisure, and home production sacrificed by taking a job. Here \( y, b, \) and \( w \) play the roles of \( u, c, \) and \( p \) in the goods market.

Emulating subsection 2.1, with \( n = N \) fixed, we get \( w = \varepsilon b + (1-\varepsilon)y \), \( V_s = \alpha_s (1-\varepsilon)(y-b) \), and \( V_b = \alpha_b \varepsilon(y-b) \). And with entry by buyers (the firms in this application), we get a similar outcome except \( n \) is endogenous and \( V_b = k_b \). With \( n = N \) fixed we have \( \partial w / \partial N \approx -\varepsilon' \), and with entry we have \( \partial n / \partial k_b < 0 \) and \( \partial w / \partial k_b \approx \varepsilon' \). If \( \varepsilon' < 0 \) then \( w \) goes up with tightness, as one might expect, but that is not true in general, as explained above for goods markets. As other features of goods markets also carry over, we proceed to applications.

Albrecht, Gautier, and Vroman (2006); Galenianos and Kircher (2009); and Kircher (2009) let workers apply for more than one job.\(^{11}\) If workers can apply to \( v \in \{1,2,\ldots\} \) vacancies, then it turns out there will be \( v \) distinct wages posted, and the optimal search strategy is to apply to one of each—that is, to look for work simultaneously in \( v \) distinct submarkets. Hence, the model exhibits wage dispersion with homogeneous agents, as is relevant because a large part of empirical wage variation cannot be explained by observables (Abowd, Kramarz, and Margolis 1999; Mortensen 2003). Also, consistent with the evidence, the density of posted wages can be shown to be decreasing, while by way of contrast, in models based on Burdett and Mortensen (1998) with homogeneous agents,

\(^{10}\)Appendix E shows \( \varepsilon'(n) \geq 0 \Leftrightarrow \sigma(n) \geq 1 \), where \( \sigma(n) \) is the elasticity of substitution. Consider a constant elasticity of substitution (CES) technology, \( \mu \)

\[
(n_1, n_2) = (n_1 + n_2)\gamma / \gamma, \gamma \in (-\infty, 1), \text{ where } \sigma = 1 / (1 - \gamma).
\]

Then \( \gamma > 0 \Rightarrow \varepsilon' > 0 \), \( \gamma < 0 \Rightarrow \varepsilon' < 0 \), and, in the Cobb–Douglas case, \( \gamma = 0 \Rightarrow \varepsilon' = 0 \). The point is that it is not hard to get \( \varepsilon' > 0 \) in examples.

\(^{11}\)One difference is: in Albrecht, Gautier, and Vroman (2006), if two or more firms make offers to the same worker they compete à la Bertrand (see also Albrecht, Gautier, and Vroman 2003; Albrecht et al. 2004); as in Galenianos and Kircher (2009) or Kircher (2009), here firms commit to posted wages.
the density is increasing. Also, again consistent with conventional wisdom, firms offering higher wages receive more applications.

Allowing multiple applications introduces an element of portfolio choice for workers, with low-wage applications serving to reduce the downside risk. This embeds in an equilibrium setting a version of Chade and Smith’s (2006) marginal improvement algorithm. For a simplified exposition, consider \( v = 2 \), so there are two wages posted, \( w_1 \) and \( w_2 \geq w_1 \), with workers sending applications to two distinct submarkets. If both pan out, they accept the highest wage; if only one pans out, they take it. Their expected payoff is therefore

\[
\begin{align*}
(8) \quad V_s &= \max_{w_1, w_2} \left\{ \alpha(n_2)(w_2 - b) + [1 - \alpha(n_2)]\alpha(n_1)(w_1 - b) \right\},
\end{align*}
\]

where \( n_j \) is the tightness in a submarket posting \( w_j \).

Generalizing the above methods, in the low-wage submarket, we solve

\[
(9) \quad V_{s1} = \max_{n_1, n_2} \alpha(n_1)(w_1 - b)
\]

subject to

\[
\frac{1 - \pi(w_1)}{n_1} \alpha(n_1)(y - w_1) = V_b
\]

where \( \pi(w) \) is the probability a worker rejects \( w_1 \) if offered. Given a solution to (9), substitute \( V_{s1} \) into (8) to obtain the problem for the high-wage submarket

\[
(10) \quad V_s = \max_{n_2, n_2} \left\{ \alpha(n_2)(w_2 - b - V_{s1}) + V_{s1} \right\}
\]

subject to

\[
\frac{\alpha(n_2)}{n_2} (y - w_2) = V_b.
\]

This looks like a problem with \( v = 1 \), but now the outside option is \( b + V_{s1} \), not just \( b \).

Since a higher outside option raises the posted wage, this is consistent with \( w_2 > w_1 \). Thus we support 2 posted wages.

For efficiency, in Galenianos and Kircher (2009), a worker who gets a high wage still enters the queue at low wages. With \( v = 2 \), if a fraction \( \rho \) of firms post \( w_1 \), then \( n_1 = \rho n_p / n_s, n_2 = (1 - \rho) n_p / n_s, \) and \( \pi(w_1) = \alpha(n_2) \). To characterize equilibrium, solve (9) and (10) with \( \rho \) set so that \( V_b \) is the same in the two submarkets. The outcome is inefficient. Since workers obtaining jobs at \( w_2 \) still enter the queue at \( w_1 \), they can prevent others from getting \( w_2 \). Neither the firms posting high wages nor the workers who obtain them take this into account, implying an unpriced externality. In Kircher (2009), workers who obtain \( w_2 \) no longer queue for \( w_1 \). This implies \( \pi(w_1) = 0 \) since any worker in the low-wage queue does not have a high-wage offer. This achieves efficiency since the unpriced externality disappears.\(^{12}\)

2.3 Summary of Baseline Models

Table 1 provides comparative statics for goods and labor markets in the benchmark model, where agents can only search in one submarket, for three cases: (i) fixed populations, (ii) entry by sellers, and (iii) entry by buyers. Most results are unambiguous, except as explained above, some of the effects on \( p \) or \( w \) can go either way. A few cases report +* or -* to indicate that the

\(^{12}\)As a referee said, it is not clear how to compare the results because one can say the environment is different if workers who obtain \( w_2 \) no longer queue for \( w_1 \). In any case, the main point here is to say that efficiency models depend on details. Wolthoff (2018) constructs a model encompassing Kircher (2009) and Galenianos and Kircher (2009), endogenizes firms’ recruitment efforts, and concludes that multiple submarkets are crucial for matching the data. Gautier and Holzner (2016) introduce a more sophisticated process to bid for workers after matching, so no vacancies remain idle because workers reject them to join firms with more applicants than they need, and that leads to efficiency. All this work constitutes progress, but there is still room for more.
signs are ambiguous, in general, but + or − in the common case $\varepsilon' \leq 0$.

The analogous table for bargaining is similar, except effects reported as $\varepsilon'$ or $-\varepsilon'$ would be 0, and those reported as $+*$ and $-*$ would be $+$ and $-$. The models above have those results in the special case of a Cobb–Douglas meeting technology, where $\varepsilon' = 0$, but in general, parameters that affect $\varepsilon$ can move trading probabilities enough to move prices in ways that are counterintuitive without understanding the theory. Under bargaining the terms of trade do not change with $n$, because while arrival rates affect expected payoffs, they do not affect the surpluses after traders meet, and hence are irrelevant in the negotiations. Now in dynamic models, as discussed below, $n$ can affect continuation values and hence the bargaining outcome, but in competitive search equilibrium $n$ affects the terms of trade even in a static environment. In any case, we highlight the following results.\[13\]

**PROPOSITION 1:** In the benchmark model with homogeneous agents that can search in at most one submarket, with or without entry, there is a unique equilibrium and it has a single price or wage. This is efficient. When agents can simultaneously search in $v > 1$ submarkets, there is a unique equilibrium and it has $v$ prices or wages. This is efficient if there are no unpriced externalities.

\[13\] As a referee pointed out, some of our propositions are really just summaries of discussions in the text. Others are more rigorous and have nontrivial proofs. At the risk of appearing pretentious, we label them all as propositions, mainly to maintain symmetry in the way we highlight key aspects of the presentation.
3. Extensions and Applications

It is desirable to consider dynamics models, where meeting probabilities translate into random durations between trades. For goods markets, we can simply repeat the static version, and since that is easy we add a few other features. Simply repeating a static model is less natural for labor, so we incorporate long-term relationships.

3.1 Goods Markets

We now embed the markets studied above into dynamic GE using the structure in Lagos and Wright (2005). Each period in discrete time, agents interact in two ways: in a decentralized market (DM) like the one analyzed above, and then in a frictionless centralized market (CM). Continue to let \( V_b \) and \( V_s \) be the DM value functions, and now let \( W_b \) and \( W_s \) be the CM value functions. In the CM, buyers solve

\[ W_b(d) = \max_{x,\ell} \{ U(x) - \ell + \beta V_b \} \]

subject to

\[ x = u\ell - d, \]

where \( \beta \) is the discount factor, \( x \) is CM numeraire, \( \ell \) is labor, \( u \) is the wage, and \( d \) is debt from the previous DM. To ease notation, assume \( x \) is produced one for one with \( \ell \), so that in equilibrium \( u = 1 \). Then the solution to (11) has \( x = x^* \) where \( U'(x^*) = 1 \). Also, the envelope condition is \( W_b'(d) = -1 \). The CM problem for sellers is omitted, but similar.

Since buyers acquire debt \( d = p \) in DM trade, we have

\[ V_b = \alpha_b[u + W_b(p)] + (1 - \alpha_b)W_b(0) \]

\[ = \alpha_b(u - p) + W_b(0) \]

because \( W_b(p) - W_b(0) = -p \), by the envelope condition. Similarly, for sellers,

\[ V_s = \alpha_s(p - c) + W_s(0). \]

Except for \( W_b(0) \) and \( W_s(0) \), notice \( V_b \) and \( V_s \) are identical to the static model. Hence, making the benchmark dynamic in this way is easy, but still nice, since higher \( \alpha(n) \) now means sellers trade faster, mapping neatly into the realm of duration analysis as used in much empirical work (e.g., Devine and Kiefer 1991). In particular, the expected times for sellers and buyers to transact are \( 1/\alpha(n) \) and \( n/\alpha(n) \).

While heterogeneity is covered more fully in section 6, it is worth seeing some examples here. Consider two types of buyers with utilities \( u_1 \) and \( u_2 > u_1 \), and entry by homogeneous sellers. Then the market segments into two distinct submarkets, \( j = 1, 2 \), where \( (p_j, n_j) \) is determined by

\[ k_s = \left[ \alpha(n_j) - n_j\alpha'(n_j) \right](u_j - c) \]

and

\[ p_j = c + k_s/\alpha(n_j). \]

There are several appealing aspects to this alternating CM–DM structure: First, we can dispense with buyers paying in the DM using transferable utility and have them pay in terms of numeraire in the next CM (and there is a genuine numeraire here, in contrast to the situation in footnote 5). Also, we can make credit imperfect to analyze debt limits and money in serious ways. Also, precisely for these reasons, this is now the workhorse model in monetary theory, and we want to use it in 4.2. Moreover, it is a tractable way to integrate search into general equilibrium, which lets one add markets for capital, other assets, etc. In fact, adding the CM actually simplifies rather dramatically the analysis of models with only DM trade. But it is not just a way to simplify things, it is realistic: some aspects of one’s actual economic life are well modeled by frictionless or centralized trade, others seem better captured by search or decentralized trade, and it seems good to have a setup with some of each. See Lagos, Rocheteau, and Wright (2017) for more.
Notice that (14) holds for \( j = 1, 2 \) independently, a feature called block recursivity. It lets us first solve for \( n_j \) in each submarket \( j \) (block 1) regardless of what is happening in other submarkets; then the number of agents in each submarket is determined (block 2) so the total number of buyers sums to the number in the economy, and free entry of sellers ensures that market tightness is correct (more on this below).

Given \( u_2 > u_1 \), one can check \( n_2 < n_1 \) and \( p_2 > p_1 \). Thus, high-valuation buyers go to submarket 2, where they pay more but trade faster. Sellers trade slower in submarket 2 and, in equilibrium, they are indifferent between it and submarket 1. This is shown in the left panel of figure 3, with buyers in submarkets 1 and 2 on indifference curves denoted \( V^*_1 \) and \( V^*_2 \), both of which are tangent to the sellers’ common indifference curve \( V^*_s = k_s \).

Now consider homogeneous buyers and two seller types in fixed numbers \( N_1 \) and \( N_2 \), with \( c_1 \) and \( c_2 > c_1 \) but the same \( k \) (Julien, Kennes, and King 2006a consider different \( k \)). Suppose \( c_j \) is not too big, so all sellers participate. As shown in the right panel of figure 3, the market segments into two submarkets where now buyers are indifferent between them. As usual, \( (p_j, n_j) \) is determined by

\[
\alpha'(n_j)(u - c_j) = V_b
\]

and

\[
p_j = \varepsilon(n_j)c_j + \left[1 - \varepsilon(n_j)\right]u.
\]

Let us normalize \( N_b = 1 \) and let \( \xi \) be the fraction of buyers in submarket 1. Then buyer indifference uniquely determines \( \xi \) by

\[
\alpha'\left(\frac{\xi}{N_1}\right)(u - c_1) = \alpha'\left(\frac{1 - \xi}{N_2}\right)(u - c_2).
\]

One can check \( n_2 < n_1 \) and \( p_2 > p_1 \), so sellers in submarket 2 trade slower, while buyers trade faster but pay higher prices.

Clearly the theory accommodates deviations from the law of one price. With heterogeneous buyers, sellers in submarket 1 settle for \( p_1 \), even though others are getting \( p_2 > p_1 \), because the latter take

---

*Figure 3. Heterogeneous Buyers (left) or Sellers (right)*
longer to sell. With heterogeneous sellers, buyers in submarket 2 pay $p_2$ even though others are paying $p_1 < p_2$, for similar reasons. This is related to, yet different from, other theories of price dispersion. In Burdett and Judd (1983), for instance, buyers see a random number of prices simultaneously—called noisy search—and when they see more than one they pick the lowest. In equilibrium ex ante identical sellers post different prices but earn equal profits, as those with lower $p$ earn less per unit but make it up on the volume. That is like our sellers, but Burdett-Judd buyers do not make a directed choice between paying less or trading faster the way they do here.

Returning to heterogeneous buyers and homogeneous sellers, consider the application to housing in Rekkas, Wright, and Zhu (2017). There are a fixed number of homogeneous houses in the market, but buyers are heterogeneous, with the value to becoming a homeowner distributed continuously across buyers with cumulative distribution function (CDF) $G(u)$ and support $[u_1, u_2]$. Now equilibrium involves a continuum of submarkets indexed by $(p_u, n_u)$ (this is treated more formally below). Hence there is a submarket for every point on sellers’ common indifference curve between $(n_1^*, p_1^*)$ and $(n_2^*, p_2^*)$, with higher $u$ associated with higher $p_u$ and lower $n_u$. Highervaluation buyers search where their trading probabilities and prices are higher, while sellers are indifferent because listing a house at a higher price means a longer average time on the market. Of course, it is no surprise that a big home in a nice neighborhood costs more than a small one in a bad neighborhood; the interest here is in residual price dispersion, the way labor economists are interested in residual wage dispersion.

What may be less obvious is that the model is consistent with sticky prices. If market conditions change, the distribution reacts, but if the change is not too big the old and new supports overlap, and sellers with $p$ in the overlapping range have no incentive to reprice. If demand falls, for example, the distribution shifts left, but many sellers can keep the same $p$. This is relevant because people claim house prices are sticky in the data and find it puzzling: “conventional wisdom is that traditional, rational, forward-looking economic theories are unable to explain extreme price stickiness of this sort, unless there are large menu costs” (Merlo, Ortaló-Magné, and Rust 2015). More generally, directed search is natural for understanding many aspects of housing markets.

Moving from houses back to generic goods, let us now make them divisible: DM buyers get a quantity or quality $q$ in exchange for payment $p$ in the next CM. Buyers’ utility and sellers’ cost, $u(q)$ and $c(q)$, satisfy the usual properties, plus $u(0) = c(0) = 0$ and $u(\bar{q}) = c(\bar{q})$ for some $\bar{q} > 0$. The efficient $q$ solves $u'(q^*) = c'(q^*)$. One can call $\hat{p} = p/q$ the unit price, unless $q$ is unobserved quality, in which case one might still call $p$ the price. We assume both $p$ and $q$ are posted, although there are alternatives—, perhaps due to limited commitment, there may be a posted unit price $\hat{p}$, and then in a meeting $q$ is chosen unilaterally by the buyer (Peters 1984) or the seller (Gomis-Porqueras, Julien, and Wang 2017).

\[ \text{In Albrecht, Gautier, and Vroman (2016), sellers first list prices, with more attractive prices meaning more buyers show up on average (although the actual number is random, as in section 5). Each buyer can accept the listed price or make a counteroffer. If no buyers accept, the seller can accept or reject the best counteroffer. If exactly one buyer accepts, he gets the house at the listed price. If two or more accept, the seller runs an auction. This is consistent with data showing that houses can sell at, above, or below listed prices. In other applications, Diaz and Jerez (2013) build a model consistent with cyclical data. Head, Sun, and Zhou (2016) have heterogenous sellers, and highly indebted homeowners tend to list high prices and take longer to sell. Hedlund (2016) has heterogenous buyers and sellers and accounts for cyclical dynamics. See also Garriga and Hedlund (2016); Stacey (2016); Moen, Nenov, and Sniekers (2019); and Head, Lloyd-Ellis, and Stacey (2017).} \]
We also introduce a limit on how much one can promise to pay, $p \leq L$, a debt/liquidity constraint that is exogenous for now, but endogenized in section 4. If it is slack then, ignoring the constants $W_b(0)$ and $W_s(0)$, we have

$$\text{(15)} \quad V_b = \max_{p,q,n} \frac{\alpha(n)}{n} \left[ u(q) - p \right]$$

subject to

$$\alpha(n) [p - c(q)] = V_s.$$ 

Indeed, when $p \leq L$ is slack, the solution has $q = q^*$, so the problem is basically the same as the one with a fixed $q$, and the usual procedure yields $(n^*, p^*)$. The generalization of (3) is $p^* = \varepsilon(n^*)c(q^*) + [1 - \varepsilon(n^*)]u(q^*)$, and the constraint is indeed slack if and only if $L \geq p^*$.

When $L < p^*$, so the constraint binds, the results are quite different. In appendix A we solve (15) and show the SOCs hold at any solution to the FOCs, so there is a unique interior solution, and it implies $L = g(q,n)$ where

$$g(q,n) = \frac{\varepsilon(n)u'(q)c(q) + [1 - \varepsilon(n)]c'(q)u(q)}{\varepsilon(n)u'(q) + [1 - \varepsilon(n)]c'(q)}.$$

This condition appears in many models with liquidity considerations and Nash bargaining (see section 4), except $\varepsilon(n)$ replaces buyers’ bargaining share $\theta$. More complicated versions of this setup are studied by Rocheteau and Wright (2005), Menzio, Shi, and Sun (2013), and Choi (2015). We again highlight the main results as follows:

**PROPOSITION 2:** Dynamic models with credit yield results analogous to static models with transferable utility, with $q$ endogenous and $p$ fixed, or vice versa. Heterogeneity segments submarkets by probability and price. Price stickiness emerges as follows: when market conditions change, some prices can stay the same, as endogenous trading probabilities make some agents indifferent to changing posted terms.

### 3.2 Labor Markets

While enduring relationships may be relevant in goods markets—one may have a favorite shop or bar—in labor markets they are ubiquitous. We now work through Moen (1997), a directed search version of Pissarides (2000), where market tightness is $n = v/(1 - e)$, the measure of vacancies over unemployment, and $e$ is the employment rate with a population of households normalized to 1.

Here we use continuous time.\(^\text{17}\) Then letting $V_{b1}$ and $V_{b0}$ be firms’ payoffs to having a worker and an open vacancy, in steady state we have

$$\text{(17)} \quad rV_{b0} = -k + \frac{\alpha(n)}{n}(V_{b1} - V_{b0}),$$

$$\text{(18)} \quad rV_{b1} = y - w + \delta(V_{b0} - V_{b1}),$$

where $k$ is the cost of a vacancy, $r$ the discount rate, and $\delta$ the job destruction rate (which is exogenous, but can be endogenized as in Mortensen and Pissarides 1994). In words, for example, (17) says the flow payoff to a vacancy is $-k$ plus the arrival rate of workers, $\alpha(n)/n$, times the gain to filling the position, $V_{b1} - V_{b0}$. Similarly, for households

$$\text{(19)} \quad rV_{s0} = b + \alpha(n)(V_{s1} - V_{s0}),$$

$$\text{(20)} \quad rV_{s1} = w + \delta(V_{s0} - V_{s1}).$$

Again it does not matter for results if firms or workers post, but the latter is easier.

\(^{17}\) There is no CM in this environment, but it can be interesting to add one (Berentsen, Menzio, and Wright 2011; Gomis-Porqueras, Julien, and Wang 2013; Zhang and Huangfu 2018; Dong and Xiao 2014).
since $V_{b0} = 0$ (free entry). This and (17)–(18) yield the wage

$$w = y - k(r + \delta)n/\alpha(n).$$

(21)

We can solve (19)–(20) for $V_{s0}$. Then, inserting $w$ from (21), we get the problem

$$rV_{s0} = b + \max_{n} \frac{\alpha(n)(y - b)}{r + \delta + \alpha(n)} - n(r + \delta)k.$$

The FOC implies $T(n) = 0$, where

$$T(n) \equiv \alpha'(n)(y - b) - \left[r + \delta + \alpha(n) - n\alpha'(n)\right]k,$$

and one can check $T(0) > 0 > T(\infty)$ and $T'(n) < 0$.

So there is a unique solution to $T(n) = 0$, and hence a unique equilibrium $n$. One can derive

$$\frac{\partial n}{\partial y} > 0, \quad \frac{\partial n}{\partial b} < 0, \quad \frac{\partial n}{\partial r} < 0, \quad \frac{\partial n}{\partial \delta} < 0, \quad \text{and} \quad \frac{\partial n}{\partial k} < 0.$$

(23)

The effects of $y$, $b$, and $k$ are consistent with table 1.2(c), plus there are new effects of $\delta$ and $r$, and all accord well with intuition. One can also show $\frac{\partial w}{\partial b} > 0$, $\frac{\partial w}{\partial \delta} < 0$, and $\frac{\partial w}{\partial r} < 0$, plus $\frac{\partial w}{\partial y} > 0$ and $\frac{\partial w}{\partial k} < 0$ if $\varepsilon' \leq 0$.

Appendix B shows the equilibrium outcome is the same as the solution to a planner’s problem posed without restricting attention to steady state—that is, the efficient $n$ solves $T(n) = 0$ at every date, as in Pissarides (2000). This is again block recursivity, where the measure of vacancies $\nu$ depends on $\varepsilon$, but tightness $n = \nu/(1 - \varepsilon)$ does not. Rearranging $T(n) = 0$, we get

$$k = \frac{\alpha(n)}{n} \frac{\varepsilon(n)(y - b)}{r + \delta + \alpha(n)[1 - \varepsilon(n)]},$$

(24)

which equates firms’ vacancy cost to their arrival rate times their share, $\varepsilon(n)$, of the appropriately discounted surplus $y - b$. This is the same as Pissarides (2000), except the elasticity $\varepsilon$ replaces firms’ bargaining share $\theta$. It matters: if we change labor-market policy, for instance, as long as $\varepsilon'(n) \neq 0$ the effects are different than predicted by bargaining.

Extensions allowing on-the-job search include Moen and Rosen (2004); Delacroix and Shi (2006); Garibaldi and Moen (2010); Schaal (2017); Tsuyuhara (2016); and Garibaldi, Moen, and Sommervoll (2016). Among other reasons, this is interesting because data show there are many direct job-to-job transitions (Fallick and Fleishman 2001; Christensen et al. 2005). As in Delacroix and Shi (2006), let $\kappa > 0$ be workers’ cost of search while employed, assumed small enough that at least some workers search while employed. This generates wage dispersion. Let $n(w)$ be the ratio of vacancies to job seekers in a submarket with wage $w$. The problem of a worker employed at $w$ is

$$rV_{s1}(w) = w + \delta\left[V_{s0} - V_{s1}(w)\right] + \max_{w',\Sigma} \left\{\alpha\left[n(w')\right]V_{s1}(w') - V_{s1}(w)\right\} - \kappa,$$

(25)

where $\Sigma = 1 (\Sigma = 0)$ indicates he engages in (abstains from) search, and if $\Sigma = 1$ then $w'$ is the next wage to which he directs his search.

An unemployed worker’s value function is similar to a worker employed at $w = b$, except it is assumed that the former has no search cost, so $V_{s0} = V_{s1}(b) + \kappa$. Also, workers are more selective in terms of the next targeted wage $w'$ when their current wage $w$ is higher (see appendix C). Solving
for equilibrium requires finding \( n(w) \). To begin, write
\[
rv_{b0} = -k + \frac{a[n(w)]}{n(w)}[v_{b1}(w) - v_{b0}],
\]
\[
rv_{b1}(w) = y - w + [\delta + \pi(w)][v_{b0} - v_{b1}(w)],
\]
where the only change from the baseline model is that jobs now end with an exogenous probability \( \delta \) plus the endogenous probability \( \pi(w) \) that a worker gets a better offer. Now free entry implies \( v_{b0} = 0 \), or
\[
k = \frac{\alpha[n(w)]}{n(w)} \frac{y - w}{r + \delta + \pi(w)}.
\]

Then \( \kappa > 0 \) implies there is a \( w \) such that workers employed at \( w \geq w \) naturally stop searching. For firms paying \( w \geq w \), \( \pi(w) = 0 \), and (26) identifies the \( n(w) \) that coincides with what one gets without on-the-job search.

Under the hypothetical situation that \( n(w) \) is computed this way everywhere, we can find the lowest wage at which the solution to (25) involves no search, and that identifies \( w \). Then, by way of induction, notice there is a minimum wage increment \( \Delta \) that workers require to justify search (appendix C). Hence, those employed at \( w \in [w - \Delta, w] \) only seek jobs with \( w' \geq w \), for which we have already determined \( n(w') \). Given \( w, w' \) is the unique solution to (25), denoted \( w' = \omega(w) \). Then \( \pi(w) = \alpha \circ n \circ \omega(w) \), where for any functions \( f \) and \( g \), \( f \circ g(x) \) denotes the composite \( f[g(x)] \). Knowing \( \pi(w) \), \( \forall w \in [w - \Delta, w] \), entry condition (26) yields \( n(w) \) at these wages. Repeating the procedure for \( w \in [w - 2\Delta, w - \Delta] \) yields \( \pi(w) \) and \( n(w) \) at those wages, and so on, until \( n(w) \) and \( \pi(w) \) are determined for all \( w \).

This establishes \( n(w) \) and \( \pi(w) \) \( \forall w \) without reference to the distribution of employment across \( w \), in and out of steady state, again due to block recursivity. Starting with higher unemployment, for example, lots of job seekers search for \( w = \omega(b) \), but also lots of firms post \( w = \omega(b) \), keeping \( n(w) \) as determined above. Thus, we can first solve for the value functions and decision rules (block 1), then study the evolution of \( e \) from any initial condition (block 2), and only in the second step does the distribution of employment come into play. Extensions of this insight allow tractable analysis of business cycle models where aggregate productivity \( y \) is stochastic. In these models, current \( y \) is enough to compute tightness in each submarket, say \( n(w, y) \), which is easier than it would be if \( n \) depended on the distribution of \( w \); see Shi (2009), Menzio and Shi (2010, 2011), Schaal (2017), and Li and Weng (2017).

To recap, there are \( n_w \) wages, \( w_1 < w_2 < \cdots < w_{n_w} \). The unemployed apply to \( w_1 = \omega(b) \); workers employed at \( w_1 \) apply to \( w_2 = \omega(w_1) \); and so on, until they stop searching. It can be shown that \( n_w \) decreases with search and entry costs. Also, simple wage contracts do not induce efficiency, similar to the model in subsection 2.2 with multiple applications. With on-the-job search, firms care about both recruitment and retention, and a single wage is not sufficient to balance the two. This is especially clear when all matches produce the same \( y \), which means on-the-job search is rent seeking that has a social cost but does not increase output. However, more complicated contracts that directly specify search activity, or specify transfers when workers quit, can restore efficiency (Menzio and Shi 2011).

Research on labor markets with directed search is a vibrant area. As regards business cycle fluctuations, in particular, Menzio and Moen (2010) show that the optimal wage contract with aggregate productivity shocks prescribes rigid wages for existing workers and downward rigidity for new hires. Menzio and Shi (2011) exploit block recursivity to
develop a tractable model of unemployment, vacancies, and job-to-job transitions over the cycle. In their model, transitions are driven by heterogeneity in firm–worker matches. They show the labor market’s response to aggregate shocks is large only if the quality of the match is observed after the match is created. Schaal (2017) also uses a directed search model to study business cycles, focusing on the impact of time-varying idiosyncratic shocks at the establishment level. Guo (2018) also uses directed search to study recessions in a model with endogenous schooling and heterogeneous agents.

**PROPOSITION 3:** The dynamic labor model without on-the-job search has a unique equilibrium and it is efficient. At each point in time, \( n \) solves \( T(n) = 0 \) and \( w \) solves (21). The outcome with on-the-job search is similar, except there is wage dispersion, and efficiency requires more complicated contracts.

4. **Monetary Economics**

Monetary theory has used random matching at least since Kiyotaki and Wright (1989), and that model has been recast using directed search by Corbae, Temzelides, and Wright (2002, 2003). We present the simple version in Julien, Kennes, and King (2008), with indivisible assets, then consider divisible assets.

4.1 **Indivisible Assets**

A fixed \([0, N]\) continuum of ex ante identical agents live forever in discrete time; there is no entry. Also, there are no centralized markets, as all trade is bilateral, and that is hindered by specialization: there are many types of goods, and it is never the case in a pairwise meeting that agent \( i \) consumes what \( j \) produces and vice versa, ruling out direct barter. Assumptions on limited commitment and private information rule out credit, so that assets have an essential role as media of exchange.

Equal measures of agents consume and produce each good. Everyone has the same utility \( u(q) \) for goods they consume and cost \( c(q) \) for those they produce. Goods are nonstorable. There is a storable asset that generates utility \( \rho \) each period for anyone holding it: if \( \rho > 0 \) this is a dividend as in standard asset-pricing theory (Lucas 1978); if \( \rho < 0 \) it is a storage cost (Kiyotaki and Wright 1989); and if \( \rho = 0 \) the asset is fiat money according to standard usage (Wallace 1980). Individual asset holdings are restricted to \( m \in \{0, 1\} \), so given a fixed supply \( M \in (0, N) \), \( M \) agents have \( m = 1 \) and act as buyers while \( N - M \) have \( m = 0 \) and act as sellers.

A novelty compared to the above models is that after trade the buyer becomes a seller and vice versa. Letting \( \Delta = V_b - V_s \) be the value to getting an asset and switching from seller to buyer, in steady state we have

\[
V_b = \rho + \alpha_b[u(q) - \beta \Delta] + \beta V_b, \tag{27}
\]

\[
V_s = \alpha_s[\beta \Delta - c(q)] + \beta V_s. \tag{28}
\]

As usual, \( \alpha_s = \alpha(n) \) and \( \alpha_b = \alpha(n)/n \) with \( n = M/(N - M) \) and \( \alpha(n) \) comes from a general meeting technology, although many papers use \( \mu = M(N - M)/N \), following

---

18 This environment is from Shi (1995) and Trejos and Wright (1995), but those papers use random search and symmetric bargaining. This is extended to other bargaining solutions by Rupert, Schindler, and Wright (2001) and Trejos and Wright (2016). There are also versions with posting and random search by Curtis and Wright (2004), or posting and noisy search by Burdett, Trejos, and Wright (2017). Wallace (2010) and references therein use abstract mechanism design. The first paper to use posting and directed search is Julien, Kennes, and King (2008), with extensions by Julien, Kennes, and Schein (2016) and He and Wright (2019).

Directed search plays two roles: first, the economy segments into markets trading different goods; second, each market segments into submarkets based on posted terms. Appendix A shows the FOCs for the submarket problem lead to

\[
\beta \Delta = \frac{\varepsilon(n)u'(q)c(q) + [1 - \varepsilon(n)]c'(q)u(q)}{\varepsilon(n)u'(q) + [1 - \varepsilon(n)]c'(q)}.
\]

The right-hand side (RHS), denoted \( g(q, \varepsilon(n)) \) in (16), is again the same as Nash bargaining except \( \varepsilon(n) \) replaces \( \theta \); different from subsection 3.1, instead of an exogenous limit \( L \), the value of assets and hence the ability to pay are endogenous.

To proceed, subtract (27)–(28) and solve for

\[
\beta \Delta = \frac{\rho + \alpha_b u(q) + \alpha_s c(q)}{r + \alpha_b + \alpha_s}.
\]

Then \( n = M/(\bar{N} - M) \) determines \( \varepsilon \), \( \alpha_b \), and \( \alpha_s \). A stationary monetary equilibrium (SME) is then a \( q \) equating the RHSs of (29) and (30), where \( q \in (0, \bar{q}) \), with \( u(\bar{q}) = c(\bar{q}) \), as required for voluntary trade. In the special but natural case \( \rho = 0 \), one can show there is a unique SME (see He and Wright 2019 for details).

To emphasize the interplay between directed search and monetary economics, consider the unique SME with \( \rho = 0 \), and the meeting technology commonly used in the literature, with \( \bar{N} = 1 \) and \( \mu = \bar{M}(1 - M) \). Then \( M = 1/2 \) is good for trade on the extensive margin, since it maximizes the number of buyer–seller meetings, but it does less well on the intensive margin, since it implies \( q < q^* \). Trejos and Wright (1995) show \( q < q^* \) at \( M = 1/2 \) using Nash bargaining with \( \theta = 1/2 \). In competitive search equilibrium, the result \( q < q^* \) at \( M = 1/2 \) follows without restrictions on \( \theta \) because \( \varepsilon(n) = 1/2 \) holds automatically at \( M = 1/2 \). Thus, we get similar results with fewer conditions, something that is typical of applications using competitive search theory.

As another connection between directed search and money, consider dynamics. The model sketched here has nonstationary equilibria for some parameters, where the value of the asset varies over time, as a self-fulfilling prophecy. That is also true with Nash bargaining models, but the microfoundations can be criticized in nonstationary equilibria (Coles and Wright 1998). This critique does not apply to competitive search, and so one can say it provides a more rigorous model of dynamics based on competitive search equilibrium, the result \( q < q^* \) at \( M = 1/2 \).

4.2 Divisible Assets

Now let agents hold any \( m \in \mathbb{R}_+ \) and bring back the frictionless CM convening after each DM. A nice feature of the CM in this kind of application is that it harnesses the distribution of \( m \), plus it allows one to incorporate many elements of mainstream macro in search models. Yet another is that we do not have to say whether agents are buyers or sellers depending on their current \( m \) as in subsection 4.1; instead we can have some called buyers that always want to consume but cannot produce in the DM, while others called sellers produce but do not consume. This would not work with only DM trade.

Focusing on \( \rho = 0 \) buyers' CM problem is

\[
W_b(m) = \max_{x, \ell, \hat{m}} \left\{ U(x) - \ell + \beta V_{b, +1}(\hat{m}) \right\}
\]
subject to
\[ x = w\ell + \phi(m - \hat{m}) - T, \]
where \( m \) is cash brought in, \( \hat{m} \) is cash taken out, \( \phi \) is its price in terms of numeraire \( x \), and \( T \) is a lump sum tax. Other than keeping track of time with the subscript on \( V_{b+1}(\hat{m}) \), (31) is like (11) with one exception: there, buyers get DM goods on credit due in the next CM; here, credit is infeasible because of issues with commitment and information, so buyers need assets as payment instruments. Still, as in subsection 3.1 we have \( W_{b}(m) = \phi \), so CM payoffs are linear. Sellers do not bring cash to the DM, but buyers may, and their FOC for \( \hat{m} > 0 \) is
\[ \phi = \beta V_{b+1}(\hat{m}). \]
Since \( m \) does not appear in this FOC, \( \hat{m} \) does not depend on what agents bring into the CM.

In the current CM, \( (p_{+1}, q_{+1}, n_{+1}) \) is posted for the next DM, where \( p_{+1} \) is the real value of the monetary payment. Since cash is a poor savings vehicle, buyers hold just enough so that \( \phi_{+1}\hat{m} = p_{+1} \). Again, it may seem natural to have sellers post and buyers search, but it is equivalent to assume the opposite. Ignoring constants and time subscripts, after some algebra, we have
\[
(32) \quad V_b = \max_{p,q,n} \left\{ \frac{\alpha(n)}{n} [u(q) - p] - i p \right\}
\]
subject to
\[ \alpha(n) [p - c(q)] = V_s \]
where \( i \) is a nominal interest rate defined by the Fisher equation, \( 1 + i = \phi/\beta \phi_{+1} \).

In stationary equilibrium \( \phi M \) is constant, so inflation is \( \dot{\phi}/\phi_{+1} = M_{+1}/M \). This plus the Fisher equation imply it is equivalent for monetary policy to peg the money growth, inflation, or nominal interest rate.

Problem (32) is the same as (15) but for one detail: buyers now must make an ex ante investment in liquidity, at cost \( ip_{+1} \), before going to the DM. Taking the FOC for \( q \), we get
\[
(33) \quad \frac{\ln / \alpha(n)}{} = \lambda(q)
\]
where \( \lambda(q) \equiv \frac{[u'(q) - c'(q)]}{c'(q)} \) is the liquidity premium. The FOC for \( n \) yields
\[
(34) \quad \alpha(n) \left[ 1 - \varepsilon(n) \right] \frac{u(q) - c(q)}{\alpha(n)} = V_s + \frac{i V_s n \varepsilon(n)}{\alpha(n)},
\]
and the constraint yields \( p = g(q, n) \), where \( g \) is defined in (16). With \( n = N \), (33) determines \( q \) and (34) determines \( V_s \); with entry by sellers, \( V_s = k_s \), and (33)–(34) determines \( (q, n) \) jointly, where in particular,
\[
(35) \quad k_s = \frac{\alpha(n) \left[ 1 - \varepsilon(n) \right] c'(q) [u(q) - c(q)]}{\varepsilon(n) u'(q) + \left[ 1 - \varepsilon(n) \right] c'(q)}.
\]
By way of comparison, consider the planner’s problem with entry by sellers,
\[
\max_{q,n} \left\{ \frac{\alpha(n)}{n} [u(q) - c(q)] - \frac{k_s}{n} \right\}.
\]
The FOCS are
\[
(36) \quad u'(q) = c'(q),
(37) \quad k_s = \alpha(n) \left[ 1 - \varepsilon(n) \right] [u(q) - c(q)].
\]
Clearly, (36) implies \( q = q^* \). From (33) this is the same as equilibrium if and only if \( i = 0 \), the Friedman rule. Then
This implies \( p = g(q, n) \) with \( \theta \) instead of \( \varepsilon \), plus

\[
\frac{ni}{\alpha(n)} = \frac{u'(q) - g_q(q, n)}{g_q(q, n)},
\]

(38)

\[
k_s = \frac{\alpha(n)(1 - \theta)c'(q)[u(q) - c(q)]}{\theta u'(q) + (1 - \theta)c'(q)}.
\]

(39)

Now (38) is the same as (33) when \( \theta = 1 \), and then it is the same as (36) if and only if \( i = 0 \). Intuitively, for buyers to make the efficient ex ante investment in liquidity, they need all the bargaining power in the DM; otherwise, \( q < q^* \) with Nash bargaining even at \( i = 0 \). But for sellers to make the efficient ex ante entry decision, they need \( \theta = \varepsilon(n) \), since that makes (39) the same as (35). As we cannot have both \( \theta = 1 \) and \( \theta = \varepsilon(n) \), in general, bargaining is not efficient; competitive search equilibrium is efficient even with two-sided holdup problems.\(^\text{20}\)

The model is generalized by Lagos and Rocheteau (2005) by fixing \( n = N \) but introducing endogenous search effort by buyers to try to capture the “hot potato” effect of inflation (people spend money faster when \( i \) is higher). As they show, this cannot happen with Nash bargaining. The reason is that \( i \) is effectively a tax on DM activity, making buyers bring less money in real terms, which lowers the gains from trade and leads to less search effort; hence, they end up spending their money slower rather than faster. With competitive search, however, even though higher \( i \) lowers the total DM surplus, it can shift the terms of trade in favor of buyers for some parameters, leading to more search effort and a “hot potato” effect.

Going back to efficiency, the problem described as “Friedman Meets Hosios” by Berentsen, Rocheteau, and Shi (2007) is that with Nash bargaining it is not generally possible to have \( \theta = \varepsilon(n) \) and \( \theta = 1 \), so even at \( i = 0 \) we cannot get \( (q^*, n^*) \) with Nash bargaining. While other bargaining solutions may do better (Aruoba, Rocheteau, and Waller 2007; Hu, Kennan, and Wallace 2009; Gu and Wright 2016) we still do not generally get efficiency at \( i = 0 \). In competitive search equilibrium we do. In a quantitative application, Rocheteau and Wright (2005) compare the effects of inflation with competitive search and Nash bargaining and find it makes a big difference.\(^\text{21}\)

In Faig and Huangfu (2007), market makers post terms for their submarkets to attract buyers and sellers, as discussed above. Recognizing that carrying currency is costly when \( i > 0 \), a shrewd market maker proposes this: all buyers pay \( \phi_b \) to enter his submarket, then buyers that meet sellers get the goods for free and all sellers collect \( \phi_s \) when they exit the DM. This allows agents to share in the cost of liquidity by eliminating cash in the hands of buyers who do not meet sellers, with market makers acting like the bankers in Berentsen, Camera, and Waller (2007). This is a nice example of how microstructure matters. In other work, Dong (2011) revisits Rocheteau, Rupert, and Wright (2007), which has indivisible labor, \( \ell \in \{0, 1\} \), as in Rogerson (1998). This generates unemployment, and there is a long-run Phillips effect.
curve exploitable by policy. Intuitively, inflation lowers \( q \) in the DM, that raises (lowers) \( x \) in the CM if the goods are substitutes (complements), and employment comoves with \( x \). Rocheteau et al. (2008) prove this in a bargaining model with \( \theta = 1 \); Dong (2011) uses competitive search and proves it with no such restriction, showing another advantage of competitive search.

Menzio, Shi, and Sun (2013) provide an alternative approach with no CM and hence a nondegenerate distribution, but it is still tractable due to block recursivity. Buyers select into submarkets: those with more \( \hat{m} \) prefer higher \( p \) and \( q \) so they can trade sooner; those with less \( \hat{m} \) prefer lower \( p \) and \( q \), even if it takes longer. Submarkets cater to their desires with different tightness, given entry by sellers. Since equilibrium separates buyers with different \( \hat{m} \), their choices are independent of the distribution, and thus so is tightness in any submarket. Sun and Zhou (2018) integrate elements of this and the alternating CM–DM framework presented above. Other natural applications include Rocheteau, Wright, and Xiao (2018) and Han et al. (2016). All this speaks to the usefulness of competitive search in monetary economies.

5. Deeper Foundations: Finite Markets


5.1 2 × 2 Market

Consider a market with two buyers and two sellers. Each seller can produce and each buyer wants to consume one unit of indivisible good \( q \), with the latter paying the former \( p \) using a divisible good \( x \) that enters payoffs linearly (i.e., we revert to transferable utility). The game proceeds as follows: first, sellers post prices; then, given \( p = (p_1, p_2) \), buyers decide where to go. If both buyers visit the same seller, one is chosen at random to get the good. Payoffs for buyers and sellers that trade are \( u - p \) and \( p - c \), and we impose \( p_j \in [c, u] \) without loss of generality. The strategy of buyer 1 is \( \gamma_{1j} \), the probability he goes to seller \( j \), and similarly for buyer 2.

Given buyer 2’s strategy, payoffs for buyer 1 from visiting sellers 1 and 2 are

\[
(40) \quad V_{b_{11}} = \left(1 - \gamma_{21} + \frac{\gamma_{21}}{2}\right)(u - p_1),
\]

\[
(41) \quad V_{b_{12}} = \left(\gamma_{21} + \frac{1 - \gamma_{21}}{2}\right)(u - p_2).
\]

In words, (40) says that at seller 1, buyer 1 gets served for sure if buyer 2 goes to seller 2, which happens with probability \( 1 - \gamma_{21} \), and gets served with probability \( 1/2 \) if buyer 2 goes to seller 1, which happens with probability \( \gamma_{21} \). Given \( p \), it is easy to check the best response of buyer 1 is

\[
(42) \quad \gamma_{11} = \begin{cases} 
0 & \text{if } \gamma_{21} > \Gamma(p) \\
[0, 1] & \text{if } \gamma_{21} = \Gamma(p) \\
1 & \text{if } \gamma_{21} < \Gamma(p)
\end{cases}
\]

where \( \Gamma(p) \equiv (u + p_2 - 2p_1)/(2u - p_1 - p_2) \).

For any \( p \), equilibrium in the stage 2 game between buyers is:

1. If \( p_1 \geq (u + p_2)/2 \) then \( (\gamma_{11}, \gamma_{21}) = (0, 0) \) (both buyers go to seller 2).
2. If \( p_1 \leq 2p_2 - u \) then \( (\gamma_{11}, \gamma_{21}) = (1, 1) \) (both buyers go to seller 1).
3. If \( (u + p_2)/2 > p > 2p_2 - u \) there are three possible equilibria:
   - \( (\gamma_{11}, \gamma_{21}) = (0, 1) \) (buyer 1 goes to seller 2, buyer 2 goes to seller 1);
   - \( (\gamma_{11}, \gamma_{21}) = (1, 0) \) (buyer 1 goes to seller 1, buyer 2 goes to seller 2);
   - \( \gamma_{11} = \gamma_{21} = \Gamma(p_1, p_2) \) (buyers play symmetric mixed strategies).
Given this and \( p_2 \) the payoff for seller 1 at stage 1 as a function of \( p_1 \) is:

1. If \( p_1 \geq (u + p_2)/2 \) then \( V_{s1} = 0 \) (seller 1 gets no buyers).
2. If \( p_1 \leq 2p_2 - u \) then \( V_{s1} = p_1 - c \) (seller 1 gets both buyers).
3. If \( (u + p_2)/2 > p > 2p_2 - u \) then two things can happen:
   - a pure-strategy equilibria with \( V_{s1} = p_1 - c \);
   - a mixed-strategy equilibrium with
     \[
     V_{s1} = \frac{3(u - p_2)(u - 2p_1 + p_2)(p_1 - c)}{(2u - p_1 - p_2)^2}.
     \]

Now the set of equilibria in pricing can be described as follows: one possibility is \( p_1 = p_2 = u \), \( \gamma_{11} = 1 \), and \( \gamma_{21} = 0 \) (for sure buyer 1 goes to seller 1 and buyer 2 goes to seller 2), which is an equilibrium since buyers can do no better at this \( \mathbf{p} \), and sellers can never do better than this. Symmetrically, \( p_1 = p_2 = u \), \( \gamma_{11} = 0 \), and \( \gamma_{21} = 1 \) is an equilibrium. Burdett, Shi, and Wright (2001) show there are also many asymmetric equilibria supported by triggers.\(^{22}\) Rather than dwelling on these, let us focus on symmetric mixed-strategy equilibria, as in much of the literature. This can be motivated by arguing that pure strategies rely on a lot of coordination, which may be reasonable in a \( 2 \times 2 \) market, but is less so in large markets (Bland and Loertscher 2012, Norman 2016).

Therefore, consider non-coordinated equilibria where buyers mix at stage 2. From the above discussion, this requires \( (u + p_2)/2 > p_1 > 2p_2 - u \), shown as the unshaded region in figure 4. In this region, choosing \( p_1 \) to maximize \( V_{s1} \) leads to

\[
(44) \quad p_1 = \frac{2(u^2 - p_2c) + (u - p_2)(p_2 + 2c)}{5(u - p_2) + 2(u - c)}.
\]

Now (44) defines seller 1’s best response \( p_1 = R(p_2) \), and symmetrically \( p_2 = R(p_1) \).

As figure 4 shows, there is a unique stage 1 equilibrium in the class under consideration, \( p_1 = p_2 = (u + c)/2 \). Then the unique stage 2 equilibrium has \( \gamma_{11} = \gamma_{12} = 1/2 \), which means buyers pick sellers at random (although the fact that search can be directed off the equilibrium path disciplines prices on it).

In this equilibrium, half the time both buyers visit the same seller, so one buyer and seller do not trade: as in Lagos (2000), simultaneously some buyers do not get served and some sellers have no customers. In fact, \( \mu = 3/2 \) is the expected number of trades, which is inefficient in the sense that \( \mu = 2 \) is physically feasible. Yet in another sense it is fairly good. Suppose both buyers go to seller 1 with arbitrary probability \( \gamma \). Then the chance they both go to the same seller is \( \gamma^2 + (1 - \gamma)^2 \), which is minimized at \( \gamma = 1/2 \). Also, while the total expected surplus is not maximized, buyers like this equilibrium because \( V_b = V_s = \)

\(^{22}\) Here is the idea: Pick any \( \mathbf{p} \) such that it is an equilibrium for buyer 1 to go to seller 1 and buyer 2 to go to seller 2 for sure. If any seller deviates, let buyers play this equilibrium at stage 2: If \( p_1 \geq (u + p_2)/2 \) both go to seller 2; if \( p_1 \leq 2p_2 - u \) both go to seller 1; and if \( (u + p_2)/2 > p > 2p_2 - u \) they play the mixed equilibrium. Burdett, Shi, and Wright (2001) characterize the set \( \mathcal{P} \) such that \( \mathbf{p} \in \mathcal{P} \) allows no profitable deviation. Intuitively, there is no profitable deviation that makes both buyers go to the same seller at stage 2, and it is not easy to find a profitable deviation that leads to a mixed equilibrium at stage 2, because sellers do poorly in mixed equilibria. For any \( \mathbf{p} \) such that profits are not too low, no one wants to deviate and trigger mixing at stage 2. See also Camera and Kim (2016), who study an infinitely repeated version.
3(u + c)/8, while they get \( V_b = 0 \) in pure-strategy equilibrium with \( p_1 = p_2 = u \). We summarize as follows.

**PROPOSITION 4:** In the 2 × 2 market there exists a unique non-coordinated equilibrium, which is \( p_j = (u + c)/2 \) and \( \gamma_{ij} = 1/2 \). The expected number of trades is \( \mu = 3/2 \) and the individual trading probabilities are \( \alpha_b = \alpha_s = 3/4 \).

In addition to posting, Julien, Kennes, and King (2000) consider auctions: if one buyer shows up he pays the posted price; and if two show up they bid, resulting in the Bertrand price \( \bar{p} = u \). One can think of sellers posting a reserve price, denoted \( p \) below. Given this, the analogs of (40) and (41) are

\[
V_{b11} = (1 - \gamma_{21})(u - p_1)
\]

and

\[
V_{b12} = \gamma_{21}(u - p_2),
\]

because a buyer gets 0 surplus unless he is the only one visiting a seller. A mixed-strategy equilibrium at stage 2 entails \( \gamma_{11} = \gamma_{21} \),

---

*Figure 4. Best Response in Price Posting in the 2 × 2 Market*
\[
= (u - p_1)/(2u - p_1 - p_2).
\]
Then one can check that sellers have a dominant strategy, \(p = (u + c)/2\). So in equilibrium, the reserve price is same as the posted price in the benchmark model. However, profit is higher: \(V_s = (u + c)/2\) with auctions and \(V_s = 3(u + c)/8\) with posting.

Coles and Eeckhout (2003) integrate the approaches by letting sellers post \(p\) contingent on the number of buyers that show up, say \(p^Q\) (\(Q\) for queue length). This nests pure posting with \(p^1 = p^2\), and auctions with \(p^2 = \bar{p} = u\) and \(p^1 = p\). They show there is an equilibrium with \(p^1 = p^2 = (u + c)/2\), as in the baseline model, but there are many others as well, all with \(p^1 = (u + c)/2\) and any \(p^2 \in [c, u]\). They are not payoff equivalent, and profit is highest with auctions. To see why, note first that a seller is indifferent to posting any pair \((p^1, p^2)\) delivering \(V_b\) to buyers, but not indifferent to his rival’s posting. Suppose, for example, a seller lowers his \(p\) to increase the probability that customers come to him. That decreases the probability that they go to his rival, so if they go to his rival they are more likely to get \(p^1\) and less likely to get \(p^2\). Ergo, stealing business is harder, and as a result profit is higher, when sellers post \(p^1 < p^2\) rather than \(p^1 = p^2\).

An exceptional case is the Coles–Eeckhout equilibrium where \(p^2 = c\) is at its minimum value. In this equilibrium a buyer gets the same expected payoff, \((u - c)/2\), whether or not the other one shows up, with the expectation taken before it is determined who gets the good if both show up. Hence, a seller’s deviation does not affect buyers’ expected payoffs when they visit his rival, and the strategic effect in the previous paragraph is inoperative—so one might say the market utility approach is valid. We say more on this in subsection 5.3.23

5.2 The \(n_b \times n_s\) Market

Consider any integer numbers of buyers and sellers, \(n_b\) and \(n_s\). A pure price posting strategy for seller \(j\) is \(p_j \in [c, u]\), and we let \(p = (p_1, \ldots, p_{n_s})\). A search strategy for buyer \(i\) is \(\gamma_i = (\gamma_{i1}, \ldots, \gamma_{in})\), with \(\gamma_{ij}\) the probability he visits seller \(j\), and we let \(\gamma = (\gamma_1, \ldots, \gamma_{n_b})\). An equilibrium is a \((p, \gamma)\) such that no one wants to deviate, and a symmetric equilibrium is one with \(p_j = p\) and \(\gamma_{ij} = 1/n_s\). To check for equilibrium we need to know what happens after a seller deviates. Starting at a symmetric \(p = (p, p, \ldots, p)\), suppose some seller, say \(j = 1\), deviates to \(p_1\) so that \(p = (p_1, p_{-1})\). We now check if this is profitable, given a symmetric equilibrium after the deviation.

Let the probability any buyer visits seller 1 after his deviation be \(\gamma_1 = \gamma_1(p_1, p_{-1})\). A symmetric subgame-perfect equilibrium is described by \(p\) and \(\gamma_1(p_1, p_{-1})\) satisfying the following conditions: (i) \(p_1 = p\) maximizes \(V_s1(p_1, p_{-1})\); (ii) \(\gamma_1(p_1, p_{-1})\) constitutes an equilibrium in the subgame for any \(p_1\) and \(p_{-1} = (p, \ldots, p)\); and (iii) on the equilibrium path \(\gamma_{ij} = 1/n_s\), while after a deviation buyers go to seller 1 with probability \(\gamma_1 = \gamma_1(p_1, p_{-1})\) and all other sellers with probability \(\tilde{\gamma}_1 = (1 - \gamma_1)/(n_s - 1)\).

PROPOSITION 5: In an \(n_b \times n_s\) market, let \(n = n_b \times n_s\) and \(\eta = 1 \eta = 1 - 1/n_s\). Then

23 Let us first mention a way around the Coles–Eeckhout indeterminacy, due to Dutu, Huangfu, and Julien (2011), in monetary economies. They show there is an equilibrium with \(p^1 = (u + c)/2\) and any \(p^2 \in [c, u]\) at \(i = 0\). By continuity, for \(i > 0\) there is an equilibrium with \(p^1\) close to \((u + c)/2\), but more care is needed to determine \(p^2\), as buyers must decide in the CM how much \(m\) to bring to the DM. Consider equilibrium where some bring \(\phi m = p^1\) and others \(\phi m = p^2\), and they are indifferent, taking into account the cost \(\phi m\). This provides an additional equilibrium condition to pin down \(p^2\). Hence, this particular indeterminacy vanishes in monetary economies (although introducing money as usual engenders other types of multiplicity). Again this shows how directed search and monetary economies are intimately related.
there exists a unique non-coordinated equilibrium, and in this equilibrium every seller sets

\[ p = \frac{(1 - \eta^{n_{b}} - n\eta^{n_{b}-1})u + n\eta^{n_{b}}c}{1 - \eta^{n_{b}} - n\eta^{n_{b}-1} + n\eta^{n_{b}}}, \tag{45} \]

while every buyer visits each seller with probability 1/\(n_{s}\). The expected number of trades is

\[ \mu = \mu(n_{b}, n_{s}) = n_{s}(1 - \eta^{n_{b}}). \tag{46} \]

\[ \alpha_{s_{1}} = 1 - (1 - \gamma_{1})^{n_{b}}. \tag{47} \]

Let the probability a buyer trades if he visits seller 1 be \(\alpha_{b_{1}} = \alpha_{b_{1}}(p_{1}, p_{-1})\). Notice \(n_{b}\gamma_{1}\alpha_{b_{1}} = \alpha_{s_{1}}\), as the left-hand side (LHS) is the expected number of buyers who get served by seller 1 and the RHS is his expected number of sales. This and (47) imply

\[ \alpha_{b_{1}} = \frac{1 - (1 - \gamma_{1})^{n_{b}}}{n_{b}\gamma_{1}}. \tag{48} \]

Now the profit of the deviant seller is

\[ V_{s_{1}}(p_{1}, p_{-1}) = \alpha_{s_{1}}(p_{1} - c) = [1 - (1 - \gamma_{1})^{n_{b}}](p_{1} - c), \tag{49} \]

and the payoff to a buyer visiting him is

\[ V_{b_{1}}(p_{1}, p_{-1}) = \alpha_{b_{1}}(u - p_{1}) = \frac{1 - (1 - \gamma_{1})^{n_{b}}}{n_{b}\gamma_{1}}(u - p_{1}). \tag{50} \]

Also, the payoff to a buyer visiting seller \(j \neq 1\) is

\[ V_{b_{j}}(p_{1}, p_{-1}) = \frac{1 - (1 - \gamma_{1})^{n_{b}}}{n_{b}\gamma_{1}}(u - p). \tag{51} \]

Given this, the FOC for maximizing \(\pi_{1}\) is

\[ 0 = \frac{\partial V_{s_{1}}}{\partial p_{1}} = \alpha_{s_{1}} + (p_{1} - c)\frac{\partial\alpha_{s_{1}}}{\partial p_{1}} \]

\[ = \alpha_{s_{1}} + n_{b}(1 - \gamma_{1})^{n_{b}-1}(p_{1} - c)\frac{\partial\gamma_{1}}{\partial p_{1}}. \tag{52} \]

In a symmetric mixed equilibrium in the subgame, buyers are indifferent between visiting any seller, which means \(\gamma_{1}\) satisfies

\[ \frac{1 - (1 - \gamma_{1})^{n_{b}}}{n_{b}\gamma_{1}}(u - p_{1}) = \frac{1 - (1 - \gamma_{1})^{n_{b}}}{n_{b}\gamma_{1}}(u - p). \tag{53} \]

Over the range \(\gamma_{1} \in (0, 1)\), we can derive \(\partial\gamma_{1}/\partial p_{1}\) and insert it into (52), then simplify using \(p_{1} = p\) and \(\gamma_{1} = 1/n_{s}\) to verify that a deviation is not profitable if and only if \(p\) solves (45). Hence there is a unique symmetric equilibrium where buyers mix. Galenianos and Kircher (2012) prove there are no asymmetric equilibria where buyers mix and sellers use pure strategies. All that remains is to show \(\mu\) satisfies (46), but that follows directly from (47) and (48) with \(\gamma_{1} = 1/n_{s}\). \(\blacksquare\)

Several remarks are in order. First, (45) endogenously gives \(p\) as a weighted average of \(c\) and \(u\), which might not be apparent from Burdett, Shi, and Wright (2001) because they normalize \(u = 1\) and \(c = 0\). In fact, the weight on \(u\) is the probability a seller gets at least 2 buyers, and the weight on \(c\) is the probability he gets just 1. Second, as in the \(2 \times 2\) game, in equilibrium buyers visit sellers at random, but the fact that search can be directed still disciplines prices.
Third, notice that $p$ is a smooth function of $n_b$ and $n_s$, which we think is nice. To say why, let $n_s$ be large, and note that as $n_b$ goes from $n_s - 1$ to $n_s + 1$, frictionless equilibrium theory predicts $p$ jumps from $c$ to $u$. As shown in Figure 5 with competitive search the discrete jump gets smoothed out by the frictions, which is one reason Peters (1984, 1991) and others advocated the approach in the first place.

Fourth, (46) endogenously gives $\mu(n_b, n_s)$ as an urn-ball meeting technology, used in economics at least since Butters (1977) and Hall (1979). The name reflects the fact that the number of buyers up at a seller is binomially distributed, like putting $n_b$ balls in $n_s$ urns in elementary probability theory, which converges to a Poisson distribution as $n_b, n_s \to \infty$ for fixed $n = n_b/n_s$, with $e^{-n}$ the probability a seller gets no buyers. Also note that $\mu(n_b, n_s)$ displays DRS (decreasing returns to scale) for finite $n_b$ and $n_s$, but quickly converges to CRS as $n_b$ and $n_s$ grow. Finally, while we appeal to Galenianos and Kircher (2012) to claim uniqueness with homogeneous sellers, Kim and Camera

\[ \mu_s = n_s(1 - e^{n}) \text{ if sellers post and } \mu_b = n_b(1 - e^{1/n}) \text{ if buyers post.} \]

As Shi and Delacroix (2018) point out, in this case, $\mu_s > \mu_b$ if and only if $n_s < n_b$, and hence we generate more meetings when the short side posts. They also show it does not matter when the matching function is symmetric, in the sense that the number of meetings $\mu(n_s, n_b)$ does not depend on who posts and who searches, which is a maintained assumption in many models.

\[ \text{Figure 5. Price and Market Tightness} \]
(2014) extend this to heterogeneity. These papers also consider more general meeting technologies, risk aversion, and private information.

Versions of the above results can be found in Peters (2000), Julien, Kennes, and King (2000), or Burdett, Shi, and Wright (2001). Using standard formulae, they also imply the following result.

**PROPOSITION 6:** Let \( n_b, n_s \to \infty \) holding \( n = n_b/n_s \) fixed. Then

\[
(54) \quad p(n_b, n_s) \to \left[1 - \varepsilon(n)\right]u + \varepsilon(n)c,
\]

where \( \varepsilon(n) = ne^{-n}/(1 - e^{-n}) \) is the usual elasticity, and

\[
(55) \quad \alpha_b(n_b, n_s) \to (1 - e^{-n})/n
\]

and

\[
\alpha_s(n_b, n_s) \to 1 - e^{-n}.
\]

As they did for the 2 \( \times \) 2 market, Julien, Kennes, and King (2000, 2001) also consider auctions in \( n_b \times n_s \) markets. They show that as \( n_b, n_s \to \infty \) the reserve price goes to \( p = c \), and payoffs are the same as under posting (see appendix D for details). They also consider a dynamic labor market version, which generates steady-state unemployment and wage dispersion. Also, we mention that in Julien, Kennes, and King (2000) and Burdett, Shi, and Wright (2001), with entry, equilibrium is not efficient in the \( n_b \times n_s \) case, because the reserve price is positive, and because the matching function exhibits decreasing returns to scale for finite numbers. See Julien, Kennes, and King (2005, 2006b, 2011) for a discussion on how this relates to efficiency in Mortensen (1982b).

5.3 Issues, Applications and Extensions

The above methods allow us to determine \( p \) without an exogenous bargaining solution, and to determine \( \alpha_s \) and \( \alpha_b \) without an exogenous meeting technology. While this is attractive, we keep an open mind. For the terms of trade, bargaining better captures situations without commitment and hence with holdup problems. For the trading probabilities, (55) is a special case of models where \( \alpha_s \) and \( \alpha_b \) come from a general meeting technology, which is relevant empirically to the extent that urn-ball functions can perform poorly when confronted with data (Petrongolo and Pissarides 2001). Different approaches may be appropriate in different applications.

Having said that, a clear advantage to finite numbers concerns off-equilibrium beliefs: all we need is subgame perfection. With a continuum, however, things get tricky. Given a candidate equilibrium \( p \), if one seller from a continuum deviates, what is the best response of buyers? If the deviator is measure 0 there is no response. The market utility approach skirts this issue in a way that is not entirely satisfactory. As an alternative, Galenianos and Kircher (2009) posit a set of artificial sellers with measure \( \delta \) that exogenously post every \( p \) in the relevant range. They then properly evaluate deviation payoffs for \( \delta > 0 \), and focus on equilibria that obtain when \( \delta \to 0 \). This works, but is slightly cumbersome.

An alternative going back to Montgomery (1991) is to use the market utility approach in \( n_b \times n_s \) markets by solving

\[
(56) \quad V_s = \max_{p, \gamma} \alpha_s(p - c)
\]

subject to

\[
\alpha_b(u - p) = V_b.
\]

In (56), \( \alpha_s \) is the probability a seller gets at least one buyer, \( \alpha_b \) is the probability a buyer visiting the seller gets served, and both depend on the probability buyers visit the seller \( \gamma \), as derived in subsection 5.2.
Eliminating $p$ using the constraint and taking the FOC for $\gamma$, we get $(1 - \gamma)^{n_b-1} u = V_b$. In symmetric equilibrium, all sellers post the same $p$ and $\gamma = 1/n_s$. Hence $V_b = \eta^{n_b-1} u$, where $\eta = 1 - 1/n_s$ and

$$p = \frac{(1 - \eta^{n_b} - n \eta^{n_b-1}) u + n \eta^{n_b-1} c}{1 - \eta^{n_b}}. \tag{57}$$

This is nice, but not quite right—in a $2 \times 2$ model, for instance, (45) gives $p = (u + c)/2$ while (57) gives $p = (u + 2c)/3$. To be fair, Montgomery (1991) solves the $2 \times 2$ model correctly, and acknowledges it is a “short cut” in the general case to take $V_b$ as given after a seller deviates. The difference between (45) and (57) is the presence of $-n\eta^{n_b-1} + n\eta^{n_b}$ in the denominator of the former, so the Burdett, Shi, and Wright (2001) price is higher than the Montgomery price, because it is less attractive for a seller to lower his $p$ when he recognizes this increases buyers’ utility. However, as $n_b n_s \to \infty$ holding $n$ fixed, this consideration vanishes, so Montgomery’s approach gives the right answer in large markets. In small markets, his approach can be misleading—for example, it yields efficiency in versions with entry or where sellers are heterogeneous, but only because it neglects relevant strategic effects.

Galenianos, Kircher, and Virag (2011) and Galenianos and Kircher (2012) propose a hybrid approach, solving (56) with $V_b = V_b(\gamma)$. This means sellers must offer buyers their market utility, but recognizes that it should be computed after a deviation. Given a symmetric mixed equilibrium after the deviation, and using (51) and $\gamma = (1 - \gamma)/(n_s - 1)$, we have

$$V_b(\gamma) = \frac{1 - \left(1 - \frac{1 - \gamma}{n_s - 1}\right)^{n_b}}{n_b \left(\frac{1 - \gamma}{n_s - 1}\right)} (u - p^*). \tag{58}$$

Solving (56) with this $V_b$, then imposing equilibrium, we get the $p$ in (45), which is correct for any $(n_b, n_s)$. Hence we get the right result for small markets, but the method and notation are similar to the earlier analysis of large markets.

One can extend this idea to make the analysis of finite markets closer to the continuum model. Given a finite number of buyers $N_b$ and each using strategy $\gamma_j$, $n_j = N_b \gamma_j$ is the expected queue of buyers at a seller $j$. The seller can trade if at least one buyer arrives, which has probability $1 - (1 - \gamma_j)^{N_b}$. Also, using $n_j = N_b \gamma_j$ we obtain the matching function

$$\alpha(n) = 1 - (1 - n/N_b)^{N_b}. \tag{59}$$

This satisfies usual properties, and if the numbers of buyers and sellers go to infinity holding $n$ constant, it converges to the urn-ball function, $1 - e^{-n}$.

Let $V_b(p)$ again be the payoff of buyers in the subgame after $p$ is posted. A seller $j$ who expects other sellers to set prices according to $p_{-j}$ solves

$$\max_{p_j} \alpha(n_j)(p_j - c) \tag{60}$$

subject to

$$\frac{\alpha(n_j)}{n_j}(u - p_j) = V_b(p_j, p_{-j}), \tag{61}$$

which is similar to (1), except now market utility depends on $p$. Using the constraint to eliminate $p_j$, and taking the FOC for $n_j$, we get

$$\alpha'(n_j) (u - c) - V_b(p) - n_j \frac{\partial V_b(p)}{\partial p_j} \frac{\partial p_j}{\partial n_j} = 0. \tag{62}$$

In the last term on the LHS, $\partial p_j/\partial n_j$ gives the change in $p_j$ needed to get a desired change in $n_j$, and $\partial V_b(p)/\partial p_j$ gives the change in market utility in the subgame. Note that in the continuum case, only the first two terms
in (61) appear. The last term leads to higher equilibrium prices: if a firm contemplates reducing its price, it understands that the additional customers it attracts reduce congestion at its competitors, which makes the competitors endogenously more attractive. This can lead to inefficiency if we add entry or heterogeneity.\footnote{Galenianos, Kircher, and Virag (2011) show that, with heterogeneous sellers, there is too much trade at those with high costs because the strategic effects increase prices more for those with low costs. Price ceilings, or minimum wages in labor markets, can restore efficiency.}

Norman (2016) offers another approach, based on population uncertainty (Myerson 1998, 2000). Suppose \( n_b \) and \( n_s \) are independently drawn from Poisson distributions, where sellers do not see the realization, while buyers see \( n_s \) and prices when they choose search strategies. This justifies the usual focus on symmetric equilibrium, where buyers use mixed strategies, since he shows any equilibrium is payoff equivalent to that (thus eliminating the other equilibria mentioned in footnote 22). The Poisson assumption makes the model tractable. Also, this model belongs to a general class in which prices are strategic complements, which implies there is a unique equilibrium if there is a unique symmetric equilibrium. In addition, as usual, since \( n_b, n_s \to \infty \) his outcome approaches our benchmark results. More generally, Norman’s ostensibly minor change in the environment generates many results in a tractable way, as should be useful in future applications.\footnote{He also generalizes Galenianos and Kircher (2012) and Kim and Camera (2014). First, he shows sellers’ profits are strictly concave in the relevant range. Then he shows his reduced-form game is supermodular under more general conditions than previous authors. This is useful because supermodular games have a smallest and largest equilibrium, and the existence of pure strategy equilibrium is simple. Because profit is strictly concave, mixed strategies by sellers are ruled out. He also provides a simple test for uniqueness with symmetric sellers.

On a less technical note, Burdett, Shi, and Wright (2001) consider two types of sellers with different capacity: \( n_1 \) of them have one unit for sale; \( n_2 \) have two units; and \( n_s = n_1 + 2n_2 \) is the total quantity on the market. This implies a matching function \( \mu(n_b, n_1, n_2) \), which in general does not reduce to a function of only \( (n_b, n_s) \). Intuitively, the coordination friction is worse when there are more sellers with one unit and fewer with two, holding \( n_s \) constant. In a labor application, this helps account for a changing Beveridge curve in the data, as the relationship between unemployment and total vacancies shifts with the mix of big and small firms. Lester (2010) extends this by letting firms choose how many positions to open, and studies the implications of various shocks, depending on whether job creation occurs via entry by new firms or expansion by existing firms (see also Tan 2012, Li and Tian 2013, Kultti and Mauring 2014, and Godenhielm and Kultti 2015).

In another application, Lester (2011) introduces semi-directed search: some buyers, called locals, are informed and direct their search to sellers posting favorable terms; others, called tourists, are uninformed and search randomly. He also has entry by sellers. In a large market, equilibrium can have two types of sellers: local shops that post low prices to attract the informed; and tourist shops that exploit the uninformed. However, if the fraction of informed buyers is high, no tourist shops open. He also analyzes finite markets, but only under a parameter restriction that guarantees no tourist shops open, so only one price is posted. Recently, Huang (2016) makes progress on relaxing this at least for \( n_s = 2 \). He shows there is a unique equilibrium, and it has sellers randomizing over \( p \in \{p_1, p_2, \ldots \} \subset [c, u] \).

The setup is useful for studying changes in information. Lester (2011) shows that increasing the fraction of informed buyers can increase or decrease prices, depending on parameters. This is contrary to
conventional wisdom and to several papers that say prices fall when consumers are better informed (Salop and Stiglitz 1977, Varian 1980, Burdett and Judd 1983, Stahl 1989). However, his result requires finite numbers, as the fraction of informed buyers does not affect prices with a continuum. Gomis-Porqueras, Julien, and Wang (2017) pursue this when the number of informed buyers is endogenized by costly advertising. They show more information can raise prices even in large markets. Bethune, Choi, and Wright (2020) also show more information can raise or lower prices in large markets in monetary economies. Intuitively, if buyers’ bargaining power is low at tourist shops, they get more from a marginal dollar at local shops. With better information, the fraction of local shops rises, so buyers bring more money, and then sellers raise prices.

Let us now return to Coles and Eeckhout (2003), by allowing sellers to post \( p^Q \) where \( Q \) is the number of buyers that show up. As in the \( 2 \times 2 \) market, indeterminacy obtains. However, Selcuk (2012) perturbs the environment by having risk-averse buyers, so \( V_b = EU(S^Q_b) \), where \( S^Q_b = (u - p^Q) / Q \) is buyers’ surplus when \( Q \) buyers show up. He shows that \( U'' < 0 \) eliminates the indeterminacy: there is a unique equilibrium, and it features \( p^Q = p \) \( \forall Q \). In particular, if \( U(S) = S^a \) with \( a \in (0, 1) \), in a market with \( n = n_b / n_s \) all sellers post

\[
p = \frac{(1 - e^{-n} - ne^{-n-1})u + an e^{-n}c}{1 - e^{-n} - ne^{-n-1} + an e^{-n}}.
\]

Notice \( \partial p / \partial a < 0 \), and \( p \to u \) as \( a \to 0 \). Also notice that \( a = 1 \) implies \( p \) is the same as (45) from proposition 5. Risk aversion seems an important extension to many applications of directed search, some of which are discussed below.

In fact, posting a contingent \( p^Q \) is still restrictive—why can’t buyers make or receive transfers even if they do not get served? Again, one might make assumptions to preclude this, but suppose we allow it. As in Jacquet and Tan (2012), for any mechanism in a general class, the outcome can be implemented by having a buyer who gets the good pay \( p \) and having others that show up pay \( \phi \), which can be positive or negative. If \( \phi = p - u \) then buyers are fully insured: their payoff is the same whether or not they get served. Jacquet and Tan (2012) show that when buyers are risk averse there is a unique equilibrium and it features full insurance, \( \phi = p - u \). One might say this provides strategic foundation for the market utility approach.

Returning to risk neutrality, Geromichalos (2012) explores several additional extensions in \( n_b \times n_s \) markets. First, he gives sellers the capacity to each serve up to \( \kappa \) customers, and post mechanisms announcing the number they will serve, which cannot exceed \( \kappa \) but could be less. He also allows \( p^Q \) to be contingent on the queue of buyers that show up, and allows payments from those who get served and those who do not. As in the previous paragraph, it suffices to consider mechanisms where buyers that get served pay \( p \) while buyers that show up but do not get served pay \( \phi \). He then shows that in a particular sense, this does not matter in large markets—it is payoff equivalent to have
sellers simply post one price, as in the baseline model.

To illustrate his methods, let us set capacity to $\kappa = 1$. Then, as shown in appendix F, in the limit as $n_b, n_s \to \infty$ with $n = n_b/n_s$ fixed, equilibrium satisfies

\[ (p - c)(1 - e^{-n}) + n\phi = (u - c)(1 - e^{-n} - ne^{-n}). \]

This pins down $V_s = (u - c)(1 - e^{-n} - ne^{-n})$, the same as the outcome in proposition 5. Intuitively, any combination $(p, \phi)$ satisfying (62) is payoff equivalent to the outcome when each seller simply posts $p$. However, this needs to be reconciled with Virag (2011), who shows that for any market size, and the same type of mechanism $(p, \phi)$, there is always an equilibrium where sellers extract the entire surplus and market utility is 0. Hence, it is not true that all equilibria converge to the limit in (62). The difference stems from the fact that Geromichalos (2012) excludes the possibility of infinite $\phi$: as Virag (2011) makes clear, the convergence does apply if there is a bound on $\phi$, in which case the limit in (62) is valid.

Geromichalos (2014) is a sequel on the Bertrand paradox—the idea that duopoly models often have an equilibrium where price equals marginal cost (which one might expect to require many sellers). A potential resolution discussed in the literature involves capacity constraints: if a firm cannot meet market demand, a rival charging more can still get customers. He argues that this is a special case of the idea that buyers’ payoffs can fall with realized demand at a given location, and considers three examples: capacity constraints, congestion effects, and prices that depend on $Q$ (as in Coles and Eeckhout 2003). These all resolve the paradox by making buyer payoffs at a seller’s location fall with the number of buyers, so they may not all go to a seller when his $p$ is lower. While other papers typically specify demand exogenously, in directed search it is endogenous as a function of strategic behavior.

6. Heterogeneity and Sorting

6.1 A General Framework

Consider any two-sided market with types $t_1$ and $t_2$ on each side, with distributions $N_1(t_1)$ and $N_2(t_2)$ on supports $\mathcal{T}_1$ and $\mathcal{T}_2$. This encompasses buyers and sellers with heterogeneous valuations or costs, workers and firms with heterogeneous productivity, and a general theory of partnership formation when agents have heterogeneous attributes. The type distributions are fixed for now, but can be determined by entry. Also, we focus mainly on the case where types are realized before decisions are made, they could also be realized after meetings.

Each agent from side 1 posts a mechanism $s \in \mathcal{S}$, including all the information side 2 can see. With pure price posting, $s = p$ and $\mathcal{S} = \mathbb{R}_+$, while if side 1 also states his type then $s = (p, t_1)$ and $\mathcal{S} = \mathbb{R}_+ \times \mathcal{T}_1$, where for notational convenience he can state any $t_1 \in \mathcal{T}_1$, but when types are observable we assign payoff $-\infty$ to lying (merely to reduce notation). Richer mechanisms include auctions, where agents in meetings have actions like bidding, and then an action for side $i$ is $a_i \in \mathcal{A}_i$. We assume mechanisms can be ordered, and for side 1 let $n_1(s, a_1, t_1)$ be the mass of types weakly below $t_1$ posting $s$ or lower, and who play action $a_1$ or lower if matched. With a slight abuse of notation, let $n_1(s)$ be the marginal, the mass posting $s$ or lower. Similarly, distributions for side 2 indicate where they direct their search and what they do if matched. Then the equilibrium concept is based on the theory of large games (Mas-Colell 1984), where an individual’s payoff is determined by his behavior, and the distribution of others’ behavior as described by $n_1$.
and \( n_2 \), but not on the behavior of any other individual.

As applied to competitive search by Eeckhout and Kircher (2010b) and Peters (2010), a key variable is tightness in a submarket with mechanism \( s \), determined on the equilibrium path by the ratio of side 2 to side 1 agents in this submarket. It has to be consistent with actual play. If a submarket attracts a mass of agents on both sides, \( n(s) \) captures the ratio. But even when these masses shrink to zero, as is often the case with a continuum of types, this ratio \( n(s) = \frac{dn_2(s)}{dn_1(s)} \) still remains well-defined using standard measure theory.\(^{28}\)

Other key variables are the set of types on side 1 posting any \( s \), and the set on side 2 searching for \( s \). Let \( \phi_1(t_i, a_i; s) \) be the distribution of types from side \( i \) playing action \( a_i \) in submarket \( s \). On the equilibrium path this has to be consistent with actual play \( n_i(s, a_i, t_i) \), more precisely with its conditional distribution given \( s \). The functions \( \phi_1, \phi_2 \), and \( n \) are on the equilibrium path almost everywhere uniquely determined by the trading strategies \( n_1 \) and \( n_2 \), but off equilibrium we have to make additional assumptions, as discussed below.

In sum, in addition to his own type, an agent cares about these objects in a submarket: the probability of trade, the distribution of types on the other side, and payoffs within matches. These depend on his choice of \( s \) and \( a_i \), the expected number of other traders determined by \( n \), and the distribution of types and actions of other traders determined by \( \phi \). Consider type \( t_i \) from side \( i \) in a submarket with mechanism \( s \), expecting tightness \( n \) and distribution \( \phi \). Denote his probability of trade by \( \tilde{\alpha}_i(s, a_i, t_i; n, \phi) \). His payoff in a pair \( t = (t_1, t_2) \), given \( a = (a_1, a_2) \), is \( v_i(s, a, t) \), and his payoff from not trading is \( \xi_i(s, a, t) \). The expected payoff is

\[
V_i(s, a_i, t_i; n, \phi) = \tilde{\alpha}_i(s, a_i, t_i; n, \phi)E \tilde{v}_i(s, a, t) + \left[ 1 - \tilde{\alpha}_i(s, a_i, t_i; n, \phi) \right] E \xi_i(s, a, t)
\]

where expectations are with respect to \( \phi \) on the other side of the market.

As elsewhere in this survey, off equilibrium is tied to the notion of subgame perfection for side 2. Let the market utility \( U(t_2) \) be the supremum of \( V_2(s, a_2, t_2; n, \phi) \) over \( s \) and \( a_2 \). If \( n(s) > 0 \) we impose the market utility assumption, requiring that only those agents who can attain their market utility are believed to trade.\(^{29}\)

Our setup is a signaling game for side 1 if their type is unknown, and we only require that agents believe that play is in undominated strategies.\(^{30}\)

Equilibrium is defined as a distribution of strategies—what to post and where to search—plus tightness and distributions in every submarket. To make this precise, let \( \text{supp}(n_i) \) be the support of \( n_i \), let \( s_0 \) be a fictitious mechanism with 0 utility (the outside option), and let \( S_0 = S \cup \{ s_0 \} \) be the set of mechanisms plus this.

**DEFINITION 1:** Equilibrium is a list of functions \( \langle n_1, n_2, n, \phi_1, \phi_2, U \rangle \) such that:

\(^{28}\) The ratio \( n(s) = \frac{dn_2(s)}{dn_1(s)} \) is called the Radon–Nikodym derivative. It builds on the idea that for any subset \( S' \subseteq S \) the integral \( \int_{s \in S'} n(s)dn_1(s) \) is the same as the associated number of side 2 agents searching for \( s \in S' \), \( \int_{s \in S'} dn_2(s) \). It simply extends the idea of a ratio if we have finite measures of types at a given \( s \) and extends it to environments where one obtains a mass of agents only if one integrates over markets. Since it is never optimal to post an \( s \) that attracts no one, we assume \( \text{supp}(n_i(s)) \) is a subset of \( \text{supp}(n_i(s)) \), where \( \text{supp}(n_i(s)) \) is the support of \( n_i \), and this makes the Radon–Nikodym derivative well defined.

\(^{29}\) Thus, \( n(s) > 0 \) implies the support of \( \phi_2(t_2, a_2; s) \) is non-empty and includes only \( \{ t_2, a_2 \} \in T_2 \times A_2 \) subject to \( V_2(s, a_2, t_2; n, \phi) = U(t_2) \).

\(^{30}\) Thus, the support of beliefs \( \phi_1 \) does not allow combinations \( \{ t_1, a_1; s \} \) such that \( \{ a_1; s \} \) is strictly dominated for \( t_1 \) whenever there is \( \{ a_1; s \} \) that is not dominated for some \( t_1 \).
1. (maximization) \( V_i(s, a_i, t_i; n, \phi) \geq V_i(s', a_i', t_i; n, \phi) \) for all \((s, a_i, t_i) \in \text{supp}(n_i)\) and \((s', a_i') \in S_0 \times A_i;\)

2. (consistency) \( \phi \) and \( n \) are consistent with \( n_1 \) and \( n_2 \) on their support;

3. (perfection) The market utility assumption holds and beliefs are on undominated strategies;

4. (feasibility) \( \forall T_i \subseteq T, \)
   \[
   \int_{S \times A \times T_i} dn_i(s, a_i, t_i) \leq \int_{T_i} dN_i(t_i),
   \]
   with equality if \( \max_{(s, a)} V_i(s, a, t_i; n, \phi) > 0 \) for almost all \( t_i \in T_i. \)

The first condition says that agents participate in submarkets that maximize their payoffs given their beliefs. The second says beliefs are consistent with strategies. The third captures off-equilibrium behavior. The inequality in the fourth says that the measure of agents in submarkets can never exceed the measure in the population, but could be less if some agents abstain from trade; if there are strictly positive returns it is suboptimal to abstain and the condition holds with equality. There is no entry here, but with slight modification one can add it, or can change the setup so that the type of side 2 is only drawn after the meeting.\(^{31}\)

6.2 Sorting

As a first concrete example, consider a bilateral meeting technology and price posting, as in sections 2–4, where types are known ex ante and observable, and a mechanism lists a price, the type of agent 1, and the type of agent 2 that type 1 aims to attract. In our general notation, \( s = (p, t_1, t_2) \in S = \mathbb{R}_+ \times T_1 \times T_2 \). There are no actions: in a meeting agents just trade at the posted price. This is relatively tractable since trading probabilities only depend on \( n \) and types are observed, implying the function \( \alpha_i \) can be replaced by \( \alpha_i(n) \), with \( \alpha_1(n) = \alpha(n) \) and \( \alpha_2(n) = \alpha(n)/n \), as in the baseline model. Trade yields surplus \( S(t_1, t_2) \) that depends on types, while no trade yields 0. Let \( S_i(t_1, t_2) \) be the surplus of side \( i \). One can show that side 1 agents reveal their types and are approached only by the desired types from side 2.

Hence, we can treat beliefs \( \phi_1 \) and \( \phi_2 \) as degenerate, and

\[
(64) \quad V_1[(p, t_1, t_2), t_i; n, \phi] = \alpha_1(n) [S_1(t_1, t_2) + p],
\]

while something similar holds for type 2 except \( p \) enters negatively. The maximization condition in the equilibrium definition can be rewritten such that side 1 maximizes (64) under the constraint that side 2 gets market utility \( U(t_2) \),

\[
\max_{p, n, t_2} \alpha_1(n) [S_1(t_1, t_2) + p]
\]

subject to

\[
\frac{\alpha(n)}{n} [S_2(t_1, t_2) - p] = U(t_2).
\]

Using the constraint to eliminate \( p \) we get \( \max_{n, t_2} 1/\alpha(n) S(t_1, t_2) - n U(t_2) \). The difference from earlier applications is that \( t_2 \) is an argument of \( U(t_2) \), but one can show \( U(t_2) \) is increasing, and the solution is described by the FOCs with respect to both \( n \) and \( t_2 \).

This setup has been studied in Shi (2001) for urn-ball matching and in Eeckhout and

\(^{31}\)Entry by side 1 at cost \( k(t_1) \), for example, simply requires \( V_1 = k(t_1) \) (Acemoglu and Shimer 1999a, b; Shi 2001; Eeckhout and Kircher 2010b). Drawing types of side 2 after matching means the distribution in each submarket is the unconditional distribution, and maximization means side 2 seeks the highest \( V_2 \) over possible realizations of \( t_2 \) (Peters and Severinov 1997).
Kircher (2010b) for general matching. A general treatment is provided in Chade, Eeckhout, and Smith (2017). An insight worth emphasizing is that a completely symmetric economy with positive assortative matching has a constant queue length everywhere, which may be surprising, as higher types on side 1 are more valuable and should be matched more frequently. But they trade with higher types on side 2, and this cancels the effect. The general relationship between prices and probabilities can be complex. When either the type distributions or the payoff functions are not symmetric, the queue length changes according to differential equations derived from the FOC and feasibility, which capture both the relative scarcity of types and the relative importance in value generation. Depending on which side reaps the surplus, prices may increase or decrease with type. Since both can be varied at the same time, it is easy to get higher prices associated with higher trading probabilities when types are heterogeneous.

This discussion relies on higher types trading with each other. If $S$ has CRS, that occurs if the inverse elasticity of substitution in production outweighs the elasticity of substitution in the matching function (Eeckhout and Kircher 2010b), meaning complementarities in production are stronger than supermodular, the usual condition in sorting models. The reason is that under weak supermodularity, $S = t_1 + t_2$, it is optimal to raise the trading probabilities of high types by pairing them with many agents from the other side, and the least-cost way to do so is to pair them with low types on the other side, who do not mind waiting as much.

7. Private Information with Bilateral Meetings

Next, consider private information about types, such as workers with better (or worse) knowledge than their employers about skills, or sellers with better (or worse) knowledge than buyers about their goods. We focus for now on the case where agents meet in pairs; multilateral meetings are covered in section 8.

7.1 Match-Specific Models

Faig and Jerez (2005) study static goods markets where buyers have private information about their valuation, Guerrieri (2008) analyzes labor markets where workers have private information about the disutility of work, and Moen and Rosén (2011) study steady state in labor markets where workers have private information about match quality and effort. Among other issues, they study efficiency. In the baseline models analyzed above, equilibria are typically constrained efficient, but this is less clear with private information. Faig and Jerez (2005) show that equilibrium is constrained efficient if the shock is realized after the match and not otherwise. Guerrieri (2008) shows that even in the first case equilibrium is inefficient outside the steady state (see also Guerrieri 2007).

Consider a version of Guerrieri (2008) by introducing match-specific private information in the goods market of subsection 2.1.

---

32 Davoodalhosseini (2015) adds entry and private information about 2’s type. See also Jerez (2014) and Mortensen and Wright (2002), where the latter is a dynamic model with endogenous flows into the market, but only considers separable payoffs. On multilateral meetings, see Shimer (2005); Julien, Kennes, and King (2005); and Albrecht, Gautier, and Vroman (2014).

33 The ease of these sorting conditions is due to the fact that each type on one side only meets a single type from the other side within a submarket. This has advantages for tractability, but does not allow someone to attribute output or wages either to the worker or the firm, as they are always matched perfectly (Eeckhout and Kircher 2011); but see subsection 8.1.
There is a unit mass of homogeneous sellers who can produce one unit at cost \( c \); buyers get utility \( u \) randomly drawn from \( F(u) \) with support \( \mathcal{U} \), privately observed after matching. The mass of active sellers is pinned down by entry at cost \( k \). Each seller that enters posts a contract where, due to the revelation principle, it is from the set of incentive compatible, individually rational, direct mechanisms \( \mathcal{S} \). Thus \( s : \mathcal{U} \mapsto [0,1] \times \mathbb{R}_+ \) specifies a menu of trading probabilities \( e(u) \) and transfers \( p(u) \) for a buyer reporting utility \( u \). Buyers observe posted contracts and decide where to go, leading to submarket tightness \( n(s) \). A buyer’s utility realization leads to a decision to trade or not. The ad interim utility of a buyer who matches in submarket \( s \), draws \( u \) and reports \( \hat{u} \) is \( v(s,u,\hat{u}) = e(\hat{u})[u - p(\hat{u})] \), and \( s \in \mathcal{S} \) is incentive compatible and individually rational if

\[
\forall s = [e(u),p(u)]_{u \in \mathcal{U}} \in \mathcal{S},
\]

(65) \( v(s,u,\hat{u}) \geq v(s,u,\bar{u}) \quad \forall u, \bar{u} \in \mathcal{U}, \)

(66) \( v(s,u,\hat{u}) \geq 0 \quad \forall u \in \mathcal{U}. \)

Equilibrium can be characterized by buyers’ market utility \( U \), the set of posted mechanisms \( \mathcal{S}^p \in \mathcal{S} \), and tightness \( n(s) \) defined on \( \mathcal{S} \), such that:

a) sellers’ maximization and free entry: \( \forall s = [e(u),p(u)]_{u \in \mathcal{U}} \in \mathcal{S}, \)

(67) \( \alpha [n(s)] \int_{u \in \mathcal{U}} e(u)[p(u) - c]dF(u) - k \leq 0, \)

with equality if \( s \in \mathcal{S}^p; \)

b) buyers’ maximization:

\[
\forall s = [e(u),p(u)]_{u \in \mathcal{U}} \in \mathcal{S},
\]

\[
(68) \frac{\alpha [n(s)]}{n(s)} \int_{u \in \mathcal{U}} e(u)[u - p(u)]dF(u) \leq U,
\]

and \( n(s) \geq 0 \) with complementarity slackness, where

\[
U = \max_{\alpha[s]} \frac{\alpha[n(s)]}{n(s)} \times \int_{u \in \mathcal{U}} e(u)[u - p(u)]dF(u).
\]

One can show that equilibrium can be characterized by maximizing buyer’s payoff subject to (65), (66), and entry. Appendix F reduces the dimensionality of the constraints using methods from mechanism design (e.g., Myerson 1981). Then one can show equilibrium is unique, and only one type of contract \([e(u),p(u)]\) is posted, where the nondecreasing trading probability \( e(u) \) and tightness \( n(u) \) solve

\[
U = \max_{e(\cdot),n} \frac{\alpha[n]}{n} \int_{u \in \mathcal{U}} e(u)(u - c)dF(u)
\]

subject to

\[
\alpha(n) \int_{u \in \mathcal{U}} e(u)[u - \frac{1 - F(u)}{f(u)}]dF(u) \geq k.
\]

Entry by sellers implies buyers get the entire surplus. Equilibrium maximizes the surplus, subject to buyers getting the information rents required for truthful revelation. Only buyers that draw \( u \geq \hat{u} \) trade, with \( \hat{u} \) endogenous. In this static environment equilibrium is constrained efficient; as mentioned, in a dynamic extension out of steady state, it is generically not constrained efficient.

7.2 Individual-Specific Models

Guerrieri, Shimer, and Wright (2010); Shao (2014); Chang (2018); Guerrieri and
Shimer (2014, 2018); Chen, Doyle, and Gonzalez (2015); Williams (2019); and Davoodalhosseini (2019) all consider screening problems à la Akerlof (1970). Consider Guerrieri, Shimer, and Wright (2010), a static environment with, in their language, homogenous principals and heterogeneous agents. Types are private information and have common value for principals and agents. Principals post contracts, and agents decide where to search given beliefs. Guerrieri, Shimer, and Wright (2010) show that there exists a fully separating equilibrium, and it is the unique equilibrium.35

Consider an extension of the model in subsection 2.1 with ex ante homogenous buyers and heterogeneous sellers. Buyers (side 1) can enter at cost, sellers (side 2) can produce an indivisible good at 0 cost, where a fraction produces high quality and produces low quality, with respective values to seller of type and the probability a buyer matches with a producer of high quality and does not trade he gets information to sellers before trade. If a seller does not trade he gets or , depending on quality, where , , so there are always gains from trade. Buyers in the market post , and sellers choose where to search. A contract specifies a trading probability and a transfer , but the same outcome obtains if buyers post menus \( \{(e', p')\}, (e^h, p^h)\).

Any is associated with a submarket with tightness and a fraction of type sellers, with \( \phi(l; s) + \phi(h; s) = 1 \), so the probability a seller matches is \( \alpha[n(s)] \), and the probability a buyer matches with a seller of type is \( \alpha[n(s)] \phi(t; s)/n(s) \). For \( s = (e, p) \) define

\[
v_1(s, t) = e(\bar{v}_1 - p)
\]

and

\[
v_2(s, t) = ep + (1 - e)\bar{v}_2
\]

where \( v_2(s, t) \) and \( v_1(s, t) \) are the payoffs in submarket \( s \) conditional on meetings.

In this setting equilibrium can be defined by a pair of sellers’ market utilities \([U^l, U^h]\), tightness \( n(s) \), and market composition \( \phi(t; s) \), both defined over \( S \), a CDF \( Z(s) \), and a set of posted contracts \( S^p \in S \), satisfying:

a) buyers’ maximization and free entry,

\[
\frac{n(s)}{\alpha[n(s)]} \sum_{j = (l, h)} \phi(j; s) v_1(s, t) - k \leq 0
\]

\( \forall s \in S \), with equality if \( s \in S^p \);

b) sellers’ maximization,

\[
\alpha[n(s)] v_2(s, t) + \{1 - \alpha[n(s)]\} \bar{v}_2^t \leq U^t
\]

\( \forall s \in S \) and \( t \in \{l, h\} \), with equality if \( n(s) < \infty \) and \( \phi(j; s) > 0 \), where

\[
U^t = \max_{s \in S} \alpha[n(s)] v_2(s, t) + \{1 - \alpha[n(s)]\} \bar{v}_2^t
\]

c) and feasibility,

\[
\int_{S^p} \phi(j; s) n(s) dZ(s) \leq \pi_j
\]

\( \forall j \), with equality if \( U^t > 0 \).

35This constitutes an alternative solution to the nonexistence problem in Rothschild and Stiglitz (1976); see Guerrieri, Shimer, and Wright (2010) for the intuition and related work.

36This is again a special case of the definition in subsection 6.1. The first two conditions come from maximization coupled with perfection. Notice the second one requires maximization on and off equilibrium: if a buyer expects a seller of a given type to search for a given contract, even off equilibrium, it must be weakly optimal for that seller to do so. This delivers uniqueness. In contrast to subsection 7.1, we must specify feasibility here because entry of buyers alone is not sufficient to ensure the right market tightness per type.
Guerrieri, Shimer, and Wright (2010) prove that equilibrium is fully separating and characterized by two simple problems: any seller \( t \in \{l,h\} \) chooses a contract \( s \) and faces tightness \( n \), where \((s,n)\) solves

\[
U^t = \max_{s \in \mathcal{S}, n} \left\{ \alpha(n)v_2(s,t) + \left[ 1 - \alpha(n) \right] \tilde{v}^t_2 \right\}
\]

subject to

\[
\frac{\alpha(n)}{n}v_1(s,t) \geq k
\]

and

\[
\alpha(n)v_2(s,t') + \left[ 1 - \alpha(n) \right] \tilde{v}^t_2 \leq U^{t'}
\]

for \( t' < t \).

As might be expected with adverse selection (e.g., based on Mirrlees 1971), individual rationality is binding for low-quality sellers, incentive compatibility is binding for high-quality sellers, and equilibrium features less trade in the submarket with high quality. Equilibrium is not generally efficient, although taxes can correct the inefficiencies (Davoodalhosseini 2019).

Guerrieri and Shimer (2014) study dynamic markets with heterogeneous assets. Intuitively, by selling in a “less liquid” market, sellers convey that their assets have higher value. They use the setup to analyze financial crises, fire sales, and flight to quality. Guerrieri and Shimer (2018) introduce another dimension of private information, albeit in a static setting. Namely, investors do not know the quality of assets nor the impatience of sellers. This generates multiple equilibria with interesting efficiency implications. Chang (2018) also analyzes a model with two dimensions of private information, but collapses them into one, avoiding multiple equilibria. She covers the case where assets that are more valuable to buyers do not necessarily have higher average quality, generating semi-pooling equilibria.

These papers study screening, as the uninformed post contracts. Delacroix and Shi (2013) study signaling. Davoodalhosseini (2012) has private information, where some buyers are informed and others are not (as in random search models going back to Williamson and Wright 1994). Directed search is a promising approach to information frictions in labor, asset, and other markets, and more research on both with screening and signaling would be welcome. Especially for asset markets, this could connect well with the theories of liquidity in section 4.

8. Meetings and Mechanisms

8.1 Private Information with Multilateral Meetings

We now merge individual- and match-specific private information, focusing mostly on one-sided heterogeneity. There is a distribution \( N_2(t_2) \) with support \([0, \bar{t}_2]\) on side 2, and a mass \( N_1 \) of homogeneous agents on side 1, again with side 1 called sellers and side 2 buyers. Since types only refer to buyers, we sometimes drop the subscript. We also stick to private valuations: the value from a match is \( S_2(t) = t > 0 \) for buyers and \( S_1(t) = 0 \) for sellers. A buyer’s type \( t \) is private information, and we focus on sellers using mechanisms that do not specify types, like auctions. In this environment the literature is mainly concerned with the case where payoffs can be replicated using direct revelation mechanisms: in a meeting, the buyer reports \( t \) and payoffs are delivered as a function of the report.\(^{37}\)

\(^{37}\)It is also possible for buyers to have more information than their own type, for example, which mechanisms are posted (McAfee 1993, Epstein and Peters 1999, Peters 1999).
Here is the issue: with bilateral meetings, it is standard that sellers do not need to post anything more than a price (e.g., Riley and Zeckhauser 1983); with multilateral meetings, under mild assumptions a seller should run an auction with reserve price $p$ that generally depends on the type distribution. A priori one might think it is inefficient if generally depends on the type distribution. In fact, the results do not generally deliver the good to the buyer left on the table? Extracting rents via reserve prices, since it is possible that all the buyers that show up sellers extract rents via reserve prices, since it is possible that all the buyers that show up have $t > 0$ but $t < p$. Are there gains from trade left on the table? Extracting rents via standard price posting, as in previous sections, does not resolve this because it does not generally deliver the good to the buyer with highest valuation. In fact, the results depend on how agents meet. With random matching $p > 0$ would be the seller’s revenue-maximizing choice, which reduces to the well-studied problem of a monopolist seller.

In McAfee (1993), Peters (1997a), and Peters and Severinov (1997), each buyer directs his search toward mechanisms they most prefer, with multilateral meetings and coordination frictions as in section 5. Given $n$, with probability $P_0(n) = e^{-n}$ a seller gets no buyers, with probability $P_1(n) = n e^{-n}$ he gets one, and with $l$ buyers $P_l(n)$ follows the Poisson formula (note the multilateral nature of meetings sets this apart from section 7). They show that restricting attention to standard auctions with reserve prices is without loss of generality for an individual seller, independent of the mechanisms posted by others. Hence we have sellers use auctions with reserve prices, and refer to the mechanism by $p$, with no further actions by sellers.

Using the notation from section 6, for each buyer let $a_2$ be his bid. In a first price auction, his payoff conditional on trading is $v_2(s,a_2,t) = t - a_2$, and the seller’s payoff is $v_1(s,a_2,t) = a_2$ if $a_2 \geq p$ as specified in $s$, and the payoff from no trade is 0. Now $\phi_2(t,a_2;p)$ is the probability a buyer who directs his search to $p$ has a type weakly below $t$ and bids weakly less than $a_2$. Let $\phi_2(a_2;p,T_2)$ be the corresponding marginal probability that any type approaching mechanism $p$ bids weakly below $a_2$. Equivalently, let the probability any buyer bids weakly above $a_2$ be $\tilde{\phi}_2(a_2;p,T_2)$. The adjusted queue of buyers who pay at least $p$ is $n\tilde{\phi}_2(p;p,T_2)$. The probability of trade for a buyer is given by $\hat{\alpha}_1 = 1 - P_0[n\tilde{\phi}_2(p;p,T_2)]$. For buyers the probability of trade is $0$ if $a_2 < p$; otherwise it is $\hat{\alpha}_2(p,a_2,t;n,\phi) = P_0[n\tilde{\phi}_2(p;p,T_2)]$, assuming that $\phi$ does not have a mass point at $a_2$ (see Kim and Kircher 2015).

Peters and Severinov (1997) start with finite numbers, then take the limit, as in section 5. There are two cases: (i) buyers draw value $t$ before deciding where to search, or (ii) they draw $t$ after meeting a seller. In equilibrium, in the limit as the market gets large, case (i) implies sellers post a reserve price equal to their outside option, $p = 0$. This validates the finding in McAfee (1993), who considers a finite economy but ignores market power, as discussed in section 5. Peters and Severinov (1997) conjecture that in case (ii) sellers post $p > 0$, but the logic of their argument reviewed below does not apply, and a careful analysis in Albrecht, Gautier, and Vroman (2012) shows $p = 0$ in that case as well.38 They also show this with second price auctions, for both cases (i) and (ii), in large markets, and they show that entry by sellers is efficient. It is interesting that a single price of $p = 0$ generates efficiency along two margins: entry and

38The efficiency of the reserve price is shown by Julien, Kennes, and King (2000) in large markets with homogeneous agents. With heterogeneous sellers, Julien, Kennes, and King (2005) show that a market where all sellers use efficient reserve price induces buyers to make efficient choices across sellers, even with finite numbers. Taking all strategic considerations into account, this carries over to $2 \times 2$ markets (Julien, Kennes, and King 2002), but not $n \times n$ markets (Kim and Kircher 2015).
efficiency within meetings. With a price of 0 it is never the case that all buyers at a given seller have \( t > 0 \) but \( t < p \). Even though buyers are locked in at the time of bidding, the fact that they choose where to search based on posted \( p \) dissipates sellers' monopoly power—they still earn rents due to the frictions, but the usual monopoly considerations vanish.

The main driving force behind Peters and Severinov's argument for case (i) is illustrated as follows: Suppose all sellers offer the same \( p > 0 \). Then any seller trades with probability less than 1 due to frictions. Now a deviating seller posting \( p - \varepsilon \) attracts additional buyers, those with type above \( p - \varepsilon \) for whom the original \( p \) is too high. In a large market the deviant seller's trading probability jumps, since there are many such buyers, making the deviation profitable. Hence, the only possibility in a large market is \( p = 0 \). Despite the fact that search is directed, in equilibrium all offers are identical and buyers pick sellers at random. This coincides with the planner's outcome, since the most trade occurs if buyers visit sellers at random, if the good goes to the highest type, without precluding anyone outright from the market. Interestingly, even though in case (i) there is ex ante information that can be used when selecting where to search, the efficient equilibrium outcome (random search and where the good goes to the highest bidder) does not use that information. Thus, even in the less informative case (ii) the same outcome occurs. Competitive search is efficient, independent of when types are drawn, which is especially nice because the efficient mechanism is so simple, with \( p \) not depending on the distribution of \( t \).

The randomness of search here contrasts with section 7, but this can be explained with two types: high types outbid low types and for them, random search maximizes the number of matches, due to the well-known concavity of the urn-ball matching function. Similarly, random search maximizes the total number of matches of all types. Therefore, total efficiency is achieved (see also subsection 8.2). When sellers are heterogeneous in costs and buyers know their type before searching, search is no longer random. Rather, higher-type buyers choose a higher cost cutoff seller and trade randomly with all types where costs are better. Higher buyer types have a higher cutoff, meaning negative sorting (Peters 1997a), similar to subsection 6.2, but now sorting is imperfect because a buyer not only trades with the cutoff seller but with all higher types. This corresponds to the planner's outcome where the chance of trading is increased by spreading these buyers across sellers.

The result obtains whether buyers only observe the reserve price, or also seller type, since given \( p \) they do not care about the latter. It also obtains if they cannot see \( p \), given a penalty for lying. Sellers accept the highest bid unless it is below cost, and buyers do not submit bids below cost. Effectively this is like a first-price auction with reserve price equal to cost (Julien, Kennes, and King 2005). Finally, even if there is no penalty for lying, and messages are cheap talk, there is an equilibrium where sellers truthfully reveal type (Kim and Kircher 2015). This highlights the role of cheap talk without commitment, as in housing and labor markets where advertised prices and wages may send messages but are typically not binding (Albrecht, Gautier, and Vroman 2016). Attaining full efficiency with cheap talk depends on details (Menzio 2007). With private values, the trading patterns in models

\[39\] It is important that buyers know the correct type \( t_1 \) to know where to attempt to trade and what to bid. In a cheap-talk environment sellers post only a message \( s \in T \) about their type, but there is no penalty for lying. The question is whether there is an equilibrium in which sellers send truthful messages in the first stage. Indeed it exists and attains the efficient sorting of the fully competitive model (Kim and Kircher 2015).
with private information or with observable types are identical: Shi (2002) and Shimer (2005) study settings in which sellers post a price and priority rule for each type and find that mixing types in a submarket is a robust outcome with urn-ball matching. They also study two-sided heterogeneity.

Mangin (2017) considers the relationship between workers’ endogenous share of the surplus and the endogenous production technology in a competitive search model with wages determined by auctions. Expected match output depends on tightness through a selection channel and a generalized Hosios condition for efficiency applies. Workers’ share of the surplus depends not only on tightness, and hence general economic conditions, but also depends directly on the distribution of match-specific productivities—the type distribution. In the limit as unemployment disappears, Mangin shows the firms’ share converges to the tail index of this distribution, which captures the fatness of its tails. The key parameter governing the surplus division is thus inherited from this distribution. Thus, competitive search delivers a deeper foundation for the link between factor prices and technology, which is simply assumed with competitive factor markets. Mangin and Julien (2021) show that the generalized Hosios condition applies in several other settings where the surplus depends on tightness, explaining the efficiency results in Albrecht, Gautier, and Tan (2012). Other work includes Lester, Visschers, and Woltzoff (2017), where sellers compete on posted asking prices and agents can trade immediately or wait for an auction, like on eBay, combining elements of optimal stopping models and competitive search.

8.2 Mechanism Design

There is a literature on mechanism design and auctions where competition for buyers is a nontrivial element, but less work on how the meeting process matters. In the private valuation case, with multilateral meetings, we saw competition for buyers results in a standard auction with reserve price equal to cost. This is simple, but raises concerns: why don’t sellers use two margins, one to attract buyers and one to guarantee that buyers who show up reveal their types? Why does competition stop at \( p = 0 \) price, rather than \( p < 0 \), which attracts even more buyers? Should there be an entry payment to attract buyers and a reserve price to select the best type? When does it suffice to have only an entry payment or simply post a price?

The general issue is to know how mechanism design is affected by the way agents meet. While this has not attracted much attention yet, we can review some advances in settings with ex ante private values for
buyers as introduced at the beginning of the previous subsection \((S_2(t) = t\) and \(S_1(t) = 0\), where buyer type \(t\) is distributed on \([0, 1]\)). But in contrast to this subsection, consider now a general meeting function \(P_t(n)\) instead of urn ball, following Eeckhout and Kircher (2010a); Lester, Visschers, and Wolthoff (2015); and Cai, Gautier, and Wolthoff (2017, 2019). This is natural, for example, if each buyer approaches one seller but the latter seller only has time to deal with a random subset of \(L\) buyers. If \(L = 1\), this is a bilateral meeting technology; if \(L = \infty\) it is urn-ball; and other intermediate cases may be equally plausible.

To see how this affects the mechanism choice, as in Cai, Gautier, and Wolthoff (2017, 2019), two observations are relevant: First, posting an optimal entry payment and a standard auction with a reserve price is at least as profitable as any direct revelation mechanism, so we can focus on a second price auction where buyers pay an entry fee and then bid. Second, equilibrium is constrained efficient, which means the reserve price equals cost, and the surplus added across sellers is maximized. Consider a seller facing tightness \(n\) and a cumulative distribution of buyer types \(\phi_2(t)\). The expected social surplus at this seller is

\[
(70) \quad \int_0^n \zeta(n - \phi_2(t), n) dt,
\]

where \(\zeta(\hat{n}, n) = 1 - \Sigma P_t(n) (1 - \hat{n} / n)^j\) represents the following: in a submarket with tightness \(n\) of all buyers and a tightness \(\hat{n}\) of high-type buyers (e.g., those with types strictly above some \(t\)), the term \(\zeta(\hat{n}, n)\) represents the seller’s probability of meeting one of these high-type buyers. So the social surplus equals the integral over the chance of meeting a buyer with type above \(t\) (see Cai, Gautier, and Wolthoff 2019).

It is optimal to pool all buyers into one market if \(\zeta(\hat{n}, n)\) is concave. To see this, suppose there is a unit measure of sellers in a submarket with \(2\hat{n}_a\) high-type buyers and \(2n_a\) buyers in total, giving the relevant tightness. And suppose there is another unit measure of sellers in a submarket with \(2\hat{n}_b\) high-type buyers and \(2n_b\) buyers in total. By (70), the surplus per seller is \(t\zeta(2n_j, 2n) + (\bar{t} - t)\zeta(2\hat{n}_j, 2n)\) with \(j \in \{a, b\}\), where the first term says that as long as the seller meets someone he creates at least value \(t\), and the second says that if he meets a high type he creates additional \(\bar{t} - t\). The total surplus sums over \(j\). Now contrast this with a setting in which buyers are assigned randomly. In this case we have a measure \(2\) of sellers that each generate surplus \(t\zeta(\hat{n}_a + \hat{n}_b, n_a + n_b) + (\bar{t} - t)\zeta(\hat{n}_a + \hat{n}_b, n_a + n_b)\). If \(\zeta\) is concave it is better to pool types in one market. The result holds more generally for arbitrary distributions of types. If \(\zeta\) is not globally concave, there are always type distributions for which segregation is desirable under mild conditions.

Suppose sellers can handle up to \(L\) buyers, where the resulting \(\zeta\) is not globally concave for \(L < \infty\), and the lack of concavity is more severe for small \(L\). Eeckhout and Kircher (2010a) show more constrained sellers want to segregate buyer types into different submarkets, and since they can then infer type there is no need for complicated mechanisms (see also Auster and Gottardi 2019). However, to ensure simple price posting is optimal for any distribution of buyer types we need bilateral meetings—see Cai, Gautier, and Wolthoff (2017, 2019), who also show it is optimal for a seller that attracts \(n\) to post auctions with a reserve price equal to cost and an entry fee \(\tau\). One can show \(\tau = 0\) if \(\zeta_2 = 0\) everywhere, a feature called invariance (Lester et al. 2019). When this applies, it is optimal to pool, as competition between types is handled by the mechanism rather
than meetings. In sum, the meeting process has many implications for mechanism choice, and we encourage more work on this.

9. Other Topics

9.1 The Nash Program

The quest for strategic foundations for axiomatic bargaining is dubbed the Nash program (Binmore 1987). Competitive search provides an explicit description of a market, with finite or infinite numbers, where traders end up sharing the gains from trade in a way that can be interpreted in terms of generalized Nash, with bargaining powers and threat points determined by economic conditions. In an $n_b \times n_s$ market, (45) gives $p$ as a weighted average of $c$ and $u$, where the weight on $u$ is the probability a seller gets at least 2 buyers; in particular, in a $2 \times 2$ market $p = (u + c)/2$ is consistent with the original Nash solution. This provides an alternative to showing how generalized Nash is the limit of a non-cooperative game, with bargaining power and threat points determined by details of the game (Binmore, Rubinstein, and Wolinsky 1986).

With axiomatic bargaining the parameter $\theta$ is structural. With competitive search the surplus shares are not generally constant, but depend on economic conditions. Ignoring that can be justified only in special cases—for instance, Cobb–Douglas matching. Also, the shares generally vary across submarkets, which is important, because in markets with two-sided investments typically there is no single $\theta$ that delivers efficiency (i.e., we cannot satisfy multiple Hosios conditions with a one $\theta$). With heterogeneity, when the expected surplus depends on market tightness as in Mangin (2017) or Mangin and Julien (2021), agents’ share depends not only on tightness, but also on the type distribution. Competitive search allows agents to trade prices against both probabilities of trade and the expected surplus, ensuring agents receive a share that is commensurate with their contribution to both matching and surplus creation. Thus generating efficiency. Hence, competitive search can deliver efficiency in these situations, so it not only provides microfoundations for sharing, it can dominate bargaining, but of course, this typically requires commitment.

9.2 Large Firms

Most of the literature on labor search, directed or otherwise, concentrates on jobs rather than firms: each job is considered an entity unto itself without specifying how they aggregate into firms. This is ill-suited for studying firm size or growth. Davis, Faberman, and Haltiwanger (2013) show that matching efficiency varies substantially, and linearly, with a firm’s growth rate. To understand this, a promising avenue is to postulate that firms with $L$ workers produce $f(L)$, and can post $v$ vacancies at cost $k(v)$, generalizing formulations where they pay $k$ to post $v = 1$. At the level of an individual vacancy the model is unchanged, but aggregation differs, and if $f$ is strictly concave or $k$ strictly convex, firms grow to a finite endogenous size. Hawkins (2013) presents such a model with finite numbers, which is complicated; most subsequent work assumes that a firm posting $v$ vacancies gets a deterministic number of hires.

In Kaas and Kircher (2015), with a strictly convex $k(v)$, firms wanting to grow faster post both higher $w$ and $v$. The former induces more hires per vacancy, consistent with Davis, Faberman, and Haltiwanger (2013).
Also, they decentralize the planner’s problem, in contrast to many models with multi-worker bargaining (Stole and Zwiebel 1996; Smith 1999; Brügemann, Gautier, and Menzio 2019). Menzio and Moen (2010) show how incomplete contracts that specify \( w \) without employment guarantees imply wage floors that insure workers against modest, but not large, shocks. As in Kaas and Kircher (2015), they use a version of block recursivity to achieve tractability. Garibaldi and Moen (2010); Schaal (2017); and Garibaldi, Moen, and Sommervoll (2016) incorporate on-the-job search, but only if \( f \) or \( k \) are linear; the general case is still outstanding.

In related work, Eeckhout and Kircher (2018) introduce two-sided heterogeneity into a large-firm model with linear \( k \) and concave \( f \) to study sorting and wage implications. Julien, Kennes, and Ritter (2018) study large firms as production teams, related to the “island matching” model in Mortensen (2009). In any period, a firm can post one or more vacancies and may lose one or more workers, and may temporarily shut down if there is a minimum number necessary to operate profitably. The optimal size of a firm/team depends on the extent of frictions. Also, complementarity between workers in a team leads to wage dispersion, suggesting that recent emphasis on team production by human resource managers can contribute to inequality. This all suggests that endogenous firm size is an interesting area for future research.

9.3 Evidence

There are several approaches to assessing directed search. Some use observational data (Faberman and Menzio 2018, Marinescu and Wolthoff 2020, Banfi and Villena-Roldán 2019), while others use data from field experiments (Dal Bo, Finan, and Rossi 2013; Belot, Kircher, and Muller 2016), to see if higher wages attract more or better job applicants. This seems to be the case if one takes care that jobs are sufficiently comparable. Laboratory experiments have been used to establish that participants indeed seem to randomize as predicted by the directed search micro-foundations (Cason and Noussair 2007; Anbarci and Feltovich 2013; Helland, Moen, and Preugschat 2017; Kloosterman 2016).

Other work explores how the current wage affects the future wages when searching on the job (Gødoy and Moen 2011) or how unemployment insurance affects search (Braun et al. 2016), which seems to indicate spillovers consistent with directed search. Other work focuses on how wages differ across employment spells (Li, Peters, and Xu 2013, 2015).

The jury is still out on which dimensions the directed search approach might fail, in general, or fail relative to random search. Engelhardt and Rupert (2017), in a calibration exercise, compares competitive search to random search and largely rejects the implications of a simple competitive search model, although they fail to reject directed search with heterogeneous workers. Calibrations that use directed search without a clear comparison to random search abound, and their success seems to indicate that this model is able to address a variety of relevant issues.

9.4 Miscellany

Lester, Rocheteau, and Weill (2015) study over-the-counter financial markets in a version of Duffie, Garleanu, and Pedersen (2005) with directed search. Watanabe (2010, 2020) considers directed search in a model of middlemen with large inventories, so they are less likely to stock out; in particular, if two firms were to merge, then if one gets two customers and the other zero, and if they can share them, it would a profitable venture. A related paper is Gautier, Ho, and Watanabe (2016), who use directed search to analyze two types of middleman, those who hold inventories they get from sellers to retrade to buyers.
and those who offer platforms for buyers and sellers to trade.

Gonzalez and Shi (2010) study differences in worker arrival rates that are unknown but learned over time. This pioneers new ways to analyze dynamic markets. Peters (1991, 1994b) studies nonstationary markets where \( n \) changes over time as agents drop out after trade. Chade, Lewis, and Smith (2014) and Nagypál (2004) develop equilibrium models of directed college choice where applicants can simultaneously send out many applications. Cheremukhin, Restrepo-Echavarria, and Tutino (2014) use rational inattention to get something between pure directed and random search (although we suggest an approach closer to Burdett and Judd 1983 might be easier and more natural). Acemoglu and Shimer (2000b) let agents choose how many offers to sample before directing their search, while Bethune, Choi, and Wright (2020) let them similarly invest in information.

Several papers study risk aversion and unemployment insurance, including Acemoglu and Shimer (1999a); Rudanko (2009, 2011); Golosov, Maziero, and Menzio (2013); and Geromichalos (2015). In Geromichalos (2015), workers that do not match receive \( b \) from the government, financed by a lump sum tax. This makes it cheap for firms to be aggressive in posting, as their competitors all contribute to the unemployment caused by attracting more workers than they hire. In equilibrium, wages are too high and entry is too low, but this can be corrected by experience rating, of if firms can post contracts as in Jacquet and Tan (2012). Golosov, Maziero, and Menzio (2013) study optimal unemployment insurance, and show it is optimal to insure workers against the risk of not getting hired, but not to redistribute across workers applying to different types of jobs. Herkenhoff (2019) and Herkenhoff, Phillips, and Cohen-Cole (2016) explore with risk aversion how extended credit access can provide insurance akin to unemployment insurance, and how this affects job search, job creation, and business cycle dynamics. All these ideas are interesting and merit additional research.

10. Conclusion

This completes our tour through the theory and applications of directed search (a decision-theoretic concept) and competitive search equilibrium (a solution concept). The framework contributes to our understanding of many phenomena, like traditional search theory, but in competitive search equilibrium the terms of trade are posted and agents target counterparties. Thus, prices have an allocative role, because those posting the terms compete for customers. This internalizes search externalities, and therefore often leads to efficiency. While this general message is well established, it is not always obvious what it means. With homogeneous workers, for example, it usually suffices for firms to post a single wage, but that is not the case when workers can apply to multiple firms, or when workers are risk averse.

Directed and random search each have advantages and disadvantages. First, competitive search equilibrium may appeal to those who have philosophical issues with random matching, and the idea that traders cannot influence who they meet by their terms of trade (recall the Introduction). Also, it tends to be tractable in analyzing with otherwise difficult problems. With heterogeneity, for instance, agents may sort into submarkets, so they know in equilibrium what type they meet and that simplifies things a lot. With entry, we often get block recursivity, which is quite useful for studying, for instance, on-the-job search and firm dynamics. Moreover, since competition and efficiency emerge naturally in benchmark models, this provides a good platform for introducing complications like moral hazard.
Directed search also has disadvantages. In some versions, it looks like those who direct their search can target very specific types but are not sure to actually meet them. Another strong assumption is commitment to the posted terms of trade. Without that, random search and bargaining sometimes yield interesting and arguably realistic inefficiencies absent from competitive search models. Also, the tractability of directed search due to segmentation into homogeneous submarkets is counterfactual in some contexts, although more sophisticated models can have heterogeneity in a given submarket. Further, random search is sometimes easier because one does not have to determine submarkets endogenously. One can also say that directed search assumes agents know rather a lot, while random search assumes they know very little, and reality is somewhere in between. As mentioned, there are attempts to study intermediate cases, and this seems promising.

Competitive search is a tractable and natural way to model competition among contracts and makes it relatively straightforward to study both finite markets and limiting large markets. Again, the theory captures the idea that if you post more favorable terms potential customers come to you with higher probability, but not generally with probability 1. Agents on both sides of the market take into account prices and trading probabilities, a generalization of standard GE theory. In particular, with homogeneous buyers and sellers the price and probability of trade are the same everywhere, but the option to post different terms, and to search differently, disciplines equilibrium behavior. For these reasons and more, the framework is extremely useful, and its future seems bright. We hope this essay inspires readers to learn more about it and to continue contributing to its development.

**APPENDIX A**

We consider two scenarios. The first has indivisible \( q \), like our baseline model, but does not necessarily have perfectly transferable utility: if a buyer makes a payment \( p \) to a seller, the latter gets \( \nu(p) \) while former gets \(-\gamma(p)\), where \( \nu(0) = \gamma(0) = 0 \), \( \nu'(q) > 0 \), \( \gamma'(q) > 0 \), \( \nu''(q) \leq 0 \), and \( \gamma''(q) \geq 0 \) \( \forall q > 0 \). Transferable utility, as in subsection 2.1, is the special case, \( \nu(p) = \gamma(p) = p \). The generalization of problem (1) is

\[
V_s = \max_{p,n} \alpha(n)[\nu(p) - c] \quad \text{subject to} \quad \frac{\alpha(n)}{n}[u - \gamma(p)] = V_b.
\]

Form the Lagrangian

\[
\mathcal{L} = \alpha(n)[\nu(p) - c] + \Lambda\left\{\frac{\alpha(n)}{n}[u - \gamma(p)] - V_b\right\}.
\]

The FOCs are:

\[
\mathcal{L}_n = \alpha'(\nu - c) + \Lambda\left\{\frac{\alpha(n)}{n}(u - \gamma) - \alpha\right\} = 0,
\]

\[
\mathcal{L}_p = \alpha\nu' - \frac{\Lambda\alpha\gamma'}{n^2} = 0,
\]

\[
\mathcal{L}_\Lambda = \frac{\alpha(n)}{n}(u - \gamma) - V_b = 0.
\]
Now \( \mathcal{L}_p = 0 \Rightarrow \Lambda = n\nu'/\gamma' \) and \( \mathcal{L}_n = 0 \Rightarrow \varepsilon(\nu - c)\gamma' = (1 - \varepsilon)(u - \gamma)\nu' \), like Nash bargaining except \( \varepsilon = n\alpha'(n)/\alpha(n) \) replaces \( \theta \). At any solution to the FOCs, the bordered Hessian is

\[
H = \begin{bmatrix}
\frac{\alpha''(u - \gamma)\nu'}{\varepsilon\gamma'} + \frac{2\alpha(1 - \varepsilon)(u - \gamma)\nu'}{n^2\gamma'} & \frac{\alpha\nu'}{n} & -\frac{\alpha(1 - \varepsilon)(u - \gamma)}{n^2} \\
-\frac{\alpha\nu'}{n} & \frac{\alpha(\gamma''\nu' - \nu'\gamma')}{\gamma'} & -\frac{\alpha\gamma'}{n} \\
-\frac{(1 - \varepsilon)(u - \gamma)}{n^2} & -\frac{\alpha\gamma'}{n} & 0
\end{bmatrix}
\]

and its determinant, after simplification, is

\[
|H| = -\left(\frac{\alpha}{n}\right)^2(u - \gamma)\left[\frac{\alpha\nu'\gamma'}{\varepsilon} + \frac{\alpha(1 - \varepsilon)^2(u - \gamma)(\gamma''\nu' - \nu'\gamma')}{n^2\gamma'}\right]
\]

Standard assumptions on \( \alpha, \nu, \) and \( \gamma \) imply \( |H| > 0 \), so the SOC\( s \) hold. As a special case this applies to \( \nu(p) = \gamma(p) = p \). Also, the same method applies to the dual problem (4).

Now consider divisible goods with \( p \) fixed, as in subsection 3.1’s credit model with a binding constraint, where \( p = L \), or subsection 4.1’s monetary model with indivisible assets, where \( p = \beta\Delta \). Form the Lagrangian

\[
\mathcal{L} = \alpha(n)[p - c(q)] + \Lambda\left\{\frac{\alpha(n)}{n}[u(q) - p] - V_b\right\}
\]

With \( p \) fixed, the FOCs are:

\[
\mathcal{L}_n = \alpha'(p - c) + \frac{\Lambda(u - p)(n\alpha' - \alpha)}{n^2} = 0,
\]

\[
\mathcal{L}_q = -\alpha c' + \frac{\Lambda\alpha}{n}u' = 0,
\]

\[
\mathcal{L}_\Lambda = \frac{\alpha(n)}{n}(u - p) - V_b = 0.
\]

Now \( \mathcal{L}_q = 0 \) implies \( \Lambda = nc'/u' \). Then \( \mathcal{L}_n = 0 \) implies \( \varepsilon(p - c)u' = (1 - \varepsilon)(u - p)c' \), or \( p = g(q, n) \) with \( g(q, n) \) given by (16), again like generalized Nash except \( \varepsilon \) replaces \( \theta \). At any solution to the FOCs, the bordered Hessian is

\[
H = \begin{bmatrix}
\frac{\alpha''(u - \gamma)\nu'}{\varepsilon u} + \frac{2\alpha(1 - \varepsilon)(u - \gamma)\nu'}{n^2u'} & -\frac{\alpha c'}{n} & -\frac{\alpha(1 - \varepsilon)(u - \gamma)}{n^2} \\
-\frac{\alpha c'}{n} & \frac{\alpha(c'u'' - u'e')}{u'} & \frac{\alpha u'}{n} \\
-\frac{(1 - \varepsilon)(u - \gamma)}{n^2} & \frac{\alpha u'}{n} & 0
\end{bmatrix}
\]
and its determinant is

\[ |H| = -\left(\frac{\alpha}{n}\right)^2 (u - p) \left[ \frac{\alpha'c'u'}{\varepsilon} + \frac{\alpha(1 - \varepsilon)(u - p)(c'u'' - u'c'\varepsilon)}{n^2u'} \right]. \]

Standard assumptions imply \(|H| > 0\), so again the SOCs hold.

**APPENDIX B**

Recall the continuous-time planner problem in subsection 3.2. Normalizing the measure of households to \(N_s = 1\), we let the state variable be employment \(e\), and let the control be the measure of vacancies posted \(v\). The law of motion is \(\dot{e} = (1 - e)\alpha(n) - \delta e\), where \(n = v/(1 - e)\). Denote the value function by \(J(e)\) and write the problem as

\[
rf(e) = \max \left\{ ey + (1 - e)b - vk + J'(e) \left[ (1 - e)\alpha\left(\frac{v}{1 - e}\right) - \delta e\right] \right\}.
\]

In case it is not obvious, we derive this, following Shimer (2004). Consider the integral form of the problem

\[
J[e(t)] = \int_t^\infty e^{-r(s-t)} \{ e(s)y - [1 - e(s)]b - vk \} ds.
\]

The objective function is household utility net of vacancy posting costs. Differentiating with respect to time, we get

\[
J'[e(t)] \dot{e}(t) = -e(t)y - [1 - e(t)]b + rf[e(t)].
\]

Using this to replace \(\dot{e}(t)\) in the objective function, we arrive at (71).

One can show the value function is linear, \(J(e) = A_0 + A_1e\). This is easiest in discrete time, as in Rogerson, Shimer, and Wright (2005), where it is easy to check the mapping analogous to (71) is a contraction, with \(J(e)\) its unique fixed point. It is also easy to check that this mapping takes linear functions into linear functions. Since the set of linear functions is closed, the fixed point \(J(e)\) is linear. To get the result in continuous time one can taking the limit of the discrete time model as the period length shrinks to 0, a standard technique in search theory (e.g., Mortensen 1986). See Wright (2013) for more details.

Therefore the FOC is

\[
k = A_1\alpha'\left(\frac{v}{1 - e}\right),
\]

which implies \(n = v/(1 - e)\) is independent of \(e\). Differentiating (71), we get

\[
rA_1 = y - b - A_1\left[ \alpha\left(\frac{v}{1 - e}\right) - \alpha'\left(\frac{v}{1 - e}\right)\frac{v}{1 - e} + \delta \right].
\]

Using (73) to eliminate \(A_1\) from (72), we arrive at \(T(n) = 0\), the steady state equilibrium condition in the text. Hence, at every point in time, the planner's \(n\) is the same as the steady state equilibrium \(n\).
APPENDIX C

Recall on-the-job search from subsection 3.2. We claim there is a $w_2$ such that workers employed at $w \geq w_2$ stop searching. To begin, for a worker employed at $w_1$ searching for $w'_1$ and another employed at $w_2 > w_1$ searching for $w'_2$, the fact that both are behaving optimally implies

$$\alpha [n(w'_2)] [V_{s1}(w'_2) - V_{s1}(w_2)] \geq \alpha [n(w'_1)] [V_{s1}(w'_1) - V_{s1}(w_2)],$$

$$\alpha [n(w'_2)] [V_{s1}(w'_2) - V_{s1}(w_1)] \leq \alpha [n(w'_1)] [V_{s1}(w'_1) - V_{s1}(w_1)].$$

Now subtraction implies

$$\alpha [n(w'_2)] [V_{s1}(w_1) - V_{s1}(w_2)] \geq \alpha [n(w'_1)] [V_{s1}(w_1) - V_{s1}(w_2)].$$

Since $V_{s1}(w)$ is strictly increasing, $V_{s1}(w_1) < V_{s1}(w_2)$ and hence $\alpha [n(w'_2)] < \alpha [n(w'_1)]$. So the worker employed at $w_2$ searches in a submarket with a higher wage and lower success rate than the worker employed at $w_1$.

We now show there is a minimum wage increment $\Delta$ workers require to justify search cost $\kappa$. The gain $V_{s1}(w_0) - V_{s1}(w)$ is bounded by $(w' - w)/(\rho + \delta)$, the difference in wages over the maximum time on a job until exogenously destroyed. A employed worker that searches needs a gain that at least makes up for the cost, $V_{s1}(w_0) - V_{s1}(w) \geq \kappa$. Hence, $\kappa \leq (w' - w)/(\rho + \delta)$, or $\Delta \geq \kappa (w' - w)$. In particular, if $w$ is high enough, a worker stops searching.

APPENDIX D

Consider auctions instead of posting in the $n_b \times n_s$ market. For seller $i$ the payoff is $u - c$ unless only one buyer shows up, in which case it is $p_i - c$, or no buyers show up, in which case it is 0. Therefore,

$$(74) \quad V_{si} = n_b \gamma_i (1 - \gamma_i)^{n_b-1} (p_i - c) + \left[ 1 - n_b \gamma_i (1 - \gamma_i)^{n_b-1} + (1 - \gamma_i)^{n_s} \right] (u - c).$$

For a buyer the payoff from visiting seller $i$ is $u - p_i$ if he is alone, and 0 otherwise. In equilibrium where buyers mix, therefore,

$$(75) \quad (1 - \gamma_i)^{n_b-1} (u - p_i) = (1 - \gamma_j)^{n_b-1} (u - p_j).$$

Suppose seller $i$ deviates to $p_i$. Given buyers mix symmetrically, $\gamma_1 + (n_s - 1) \gamma_j = 1$, and

$$(76) \quad (1 - \gamma_i)^{n_b-1} (u - p_i) = \left( 1 - \frac{1 - \gamma_i}{n_s - 1} \right)^{n_b-1} (u - p_j).$$

Implicit differentiation and simplification implies

$$\frac{\partial \gamma_i}{\partial p_i} = -\frac{(1 - \gamma)(n_s - 1)}{(n_b - 1)(u - p)n_s}.$$
around the equilibrium values of \( p_i = p \) and \( \gamma = 1/n_s \). Taking the FOC from maximizing \( V_{st} \) with respect to \( p_i \) and simplifying, we get

\[
p = \frac{(n - \frac{1}{n_s})u + (n_s - 2 + \frac{1}{n_s})c}{n + n_s - 2}.
\]

As \( n_b, n_s \to \infty \) holding \( n \) fixed, \( p \to c \) and \( \pi \to (1 - e^{-n} - ne^{-n})(u - c) \). This is the same as the payoff under posting.

**APPENDIX E**

Consider any CRS technology \( \mu(n_1, n_2) \) and let \( n = n_1/n_2 \). We claim \( \sigma(n) \geq 1 \Leftrightarrow \varepsilon'(n) \geq 0 \), where \( \varepsilon(n) = n\alpha'(n)/\alpha(n) \), and \( \sigma(n) \) is the elasticity of substitution,

\[
\sigma(n) = \frac{d n}{d(\mu_2/\mu_1)} \frac{\mu_2/\mu_1}{n}.
\]

As in standard production theory, \( \sigma \) measures the degree of complementarity between inputs. Clearly, \( \mu(n_1, n_2) = n_2 \alpha(n) = n_1 \alpha(n)/n \), which implies \( \mu_1 = \alpha'(n) \) and \( \mu_2 = \alpha(n) - \alpha'(n)n \). Therefore, \( \mu_2/\mu_1 = \alpha(n)/\alpha'(n) - n \) and \( d(\mu_2/\mu_1)/dn = -\alpha(n)\alpha''(n)/\alpha'(n)^2 \). Given this, \( \sigma(n) \) implies

\[
\sigma(n) = -\frac{\alpha'(n)^2}{\alpha(n)\alpha''(n)} \frac{\alpha(n) - \alpha'(n)n}{n\alpha'(n)} = -\frac{\alpha'(n)[1 - \varepsilon(n)]}{\alpha''(n)n}.
\]

Now algebra implies

\[
\varepsilon'(n) = \left[ \frac{\alpha''(n)n + \alpha'(n)}{\alpha(n)} - n\alpha'(n)\alpha'(n) \right] \frac{\alpha(n) - n\alpha'(n)\alpha'(n)}{\alpha(n)^2}
\]

\[= \frac{\alpha'(n)}{\alpha(n)} [1 - \varepsilon(n)] [1 - 1/\sigma(n)].\]

This proves \( \varepsilon'(n) \geq 0 \Leftrightarrow \sigma(n) \geq 1 \).

For the specification \( \mu(n_1, n_2) = n_2(1 - e^{-n_1/n_2}) \) discussed in section 5, \( \alpha(n) = 1 - e^{-n} \) and \( \varepsilon(n) = ne^{-n}/(1 - e^{-n}) \). Moreover,

\[
\sigma(n) = \frac{1 - e^{-n} - ne^{-n}}{n(1 - e^{-n})} < 1,
\]

\[
\varepsilon'(n) = \frac{e^{-n}(1 - e^{-n} - n)}{(1 - e^{-n})^2} < 0,
\]

because \( 1 - e^{-n} < n \). As discussed in section 4, another common specification, especially in monetary economics following Kiyotaki and Wright (1991, 1993), is \( \mu(n_1, n_2) = n_1n_2/(n_1 + n_2) \). This implies \( \alpha(n) = n/(1 + n) \), \( \sigma(n) = 1/2 \), \( \varepsilon(n) = 1/(1 + n) \), and \( \varepsilon'(n) = -1/(1 + n)^2 < 0 \). The CES function is \( \mu(n_1, n_2) = (n_1^a + n_2^a)^{1/a} \), where \( a \in \)
This implies \( \alpha(n) = (1 + n^a)^{1/a}, \sigma(n) = 1/(1 - a), \) and \( \varepsilon(n) = n^a/(1 + n^a). \) Clearly, \( \sigma \succcurlyeq 1 \Leftrightarrow a \succcurlyeq 0 \Leftrightarrow \varepsilon'(n) = an^{-1}/(1 + n^a)^2 \succcurlyeq 0, \) providing a simple example with \( \varepsilon' > 0. \) A special case is the Cobb–Douglas function, \( \mu(n_1, n_2) = n_1^a n_2^{1-a}, \) which implies \( \alpha(n) = n^a, \sigma(n) = 1 \) and \( \varepsilon'(n) = 0. \)

**APPENDIX F**

Standard results imply the incentive compatibility and individual rationality constraints can be rewritten

\[
(78) \quad v(t_2, t_2) = v(t_2, t_2) + \int_{t_2}^{s} e(s) ds,
\]

plus \( v(t_2, t_2) \geq 0 \) and \( e(\cdot) \) nondecreasing. Using (78), we obtain

\[
\int_{t \in \mathcal{T}} v(t_2, t_2) dN_2(t_2) = v(t_2, t_2) + \int_{t \in \mathcal{T}} \int_{t_2}^{s} e(s) ds n_2(t_2) dt_2.
\]

After integrating the last term by parts, we rewrite this as

\[
\int_{t \in \mathcal{T}} v(t_2, t_2) dN_2(t_2) = v(t_2, t_2) + \int_{t \in \mathcal{T}} \frac{1 - N_2(t_2)}{n_2(t_2)} e(t_2) dN_2(t_2).
\]

Using this and the definition of \( v(t_2, t_2), \) we rewrite \( v(t_2, t_2) \geq 0 \) as

\[
\int_{t \in \mathcal{T}} e(t) \left[ t - p(t) - \frac{1 - N_2(t_2)}{n_2(t_2)} \right] dN_2(t_2) \geq 0.
\]

Hence, the relevant problem is:

\[
U = \max_{v(\cdot) \geq 0} \frac{\alpha(n)}{n} \int_{t \in \mathcal{T}} e(t_2) \left[ t_2 - p(t_2) \right] dN_2(t_2)
\]

subject to

\[
\int_{t_2 \in \mathcal{T}} e(t_2) \left[ t_2 - p(t_2) - \frac{1 - N_2(t_2)}{n_2(t_2)} \right] dN_2(t_2) \geq 0,
\]

\[
\alpha(n) \int_{t_2 \in \mathcal{T}} e(t_2) p(t_2) dN_2(t_2) = k.
\]

Using the second constraint to eliminate \( p(t_2), \) we reduce this to

\[
\alpha(n) \int_{t_2 \in \mathcal{T}} e(t_2) p(t_2) dN_2(t_2) = k.
\]

This is the problem discussed in the text. \( \blacksquare \)
REFERENCES


955–69.
▶ Dong, Mei. 2010. “Inflation and Variety.” International


