MARKETS WITH SEARCH FRICTION AND THE DMP MODEL

Prize Lecture, December 8, 2010

by

DALE T. MORTENSEN

Northwestern University and Aarhus University

1. INTRODUCTION

What are search frictions? What role do they play in the analysis of markets? Why are they important? These are the questions I will try to answer. Along the way, I will also present a short history of the development of the ideas and relate them to the current labor market conditions in Europe and the United States.

Perhaps the best way to answer the "what" question is by example. Two years ago my wife and I decided that it was time to move from our home of 35 years to a more comfortable and convenient condo. The apartment had to be relatively large and on one floor, with a view of Lake Michigan. With these features in mind, we consulted several agents to generate a list of possibilities. We then spent considerable time searching for the right one, at least one which could meet our needs and fit our budget. Finally, a year ago we bought the apartment, which has since been remodeled. So, now we need to sell the house. How do we find a buyer willing to pay the price we are asking? All of the time and effort spent by both sides of such a transaction represents search and matching frictions.

There are other markets in which search and matching activity is costly. Although only economists think in these terms, the process of finding a life partner is costly and time consuming. Of course, courtship may have some pleasures that are not present in house hunting, but the point is that an investment in information must be made in the hope of locating a fruitful long-term relationship. Obviously, all these same features are present in the process of finding a job.

We know all that, so what is all the fuss? The presence of search and matching costs has some important general implications for how markets perform their function of allocating goods and services to producers and consumers. They serve as hindrances to the process of efficient allocation not only because they reduce the gains from trade. The costs must be borne now but the returns come only in the future. Furthermore, a rule for stopping a sequential search is needed. As the future is never known with certainty, the evaluation of the prospective benefits requires the formation of expectations.

An acceptable house, partner or job, then, is one that offers an expected stream of future benefit that has a value in excess of the option to continue to search for an even better alternative.

In the standard models of markets available prior to 1970, all of these complications were ignored. The best known model is of perfect competition. In the stark model of perfect competition, the product or service traded is standardized, all properties and attributes are known to the buyer as well as the seller by assumption. In addition, trade is centralized so that all agents are aware of a common "going price." Given these conditions, together with the preferences and costs of supplying the product, the competitive price is simply that which equates today's supply and demand. That is, the equilibrium price clears the market; the quantity that the buyers want to purchase is exactly that which sellers are willing to provide.

An example of a real market in which these conditions prevail, at least approximately, is that for commodities and contracts representing commodities such as those traded on the Chicago Board of Trade. Originally, farmers brought wheat, pork bellies, soy beans, etc. to a market. In these markets, buyers and sellers came together at a designated time and place to sell the goods at auction. As these markets developed, eventually "derivatives" were traded instead in the form of contracts that provided the means for farmers to tie down prices for crops that would be available only in the future. Over the years, the technology of trade has changed in response to advances in the ability to communicate. From its origins on the streets of Chicago, the Board of Trade moved to a building housing "trading pits" for the open-outcry exchange by brokers representing buyers and sellers. Recently, this form of trading has been supplemented and to some extent displaced by trading on electronic networks. Although the extent to which these changes in exchange technology have contributed to more efficient markets is debated, information costs associated with exchange are virtually nonexistent.

Because of the search and matching costs associated with the heterogeneity of jobs, individuals, and homes, this form of centralized trading is simply not possible in labor, housing, and marriage markets. Although advances in communications technologies have affected the operation of these markets as well, the need to gather information about the properties of the unique item traded as well as the ask and bid price is still present. The three of us have been honored as pioneers in the development of a framework designed to seek answers to the following question: How are prices and quantities at which trade takes place determined in markets in which information is imperfect in this sense, and what are the social implications of the answers to this question?

2. THE FLOWS APPROACH

The theoretical lens previously used to view the labor market was the "supply and demand" framework of neoclassical economics. As noted above, this approach assumes exchange in centralized markets in which information about the goods and services traded as well as the price is perfect. In this framework, the only explanation for unemployment and vacant jobs is that the labor market sometimes "fails to clear." This view was responsible for the conclusion that unemployment could only arise when the wage level exceeded the value that equates supply and demand. Unemployment is "involuntary" when the price is above its market clearing level. Workers are unemployed because jobs are not available at the prevailing wages, period. The only recourse is to either expand the number of jobs or somehow lower the wage. As a graduate student in the early 60s, I was taught that it was the government's responsibility, to do so, using both the tools of fiscal and monetary policy as advocated by Keynes and his followers.

In the late 60s a group of economists (Stigler, 1962; Holt and David, 1966; Phelps et al., 1970; among others) started to think about a more nuanced conception of the labor market based on observation regarding the actual experiences of individual workers over time. In normal times, the typical employment relationship lasts for several years, while unemployment spells are relatively short for most workers. Furthermore, jobs are heterogeneous, offering different wages, long-term career opportunities, and other characteristics such as job security, while workers vary with respect to skills and preferences for job characteristics other than wage. These facts suggested that workers and employers find it in their interest to invest time and effort in the process of discovering a "good" job-worker match. In this view, unemployed workers and unfilled jobs are those engaged in this process.

Early on, theorists realized that a dynamic "flows approach" was needed for an adequate analysis of unemployment fluctuations. Namely, the theory should speak to observations on the duration and incidence of unemployment and how these move in response to the fluctuations of business activity. Because investments in information are required prior to forming an acceptable relationship, the relationships that do form are relatively long-lived. As an implication of these observations, the number of existing relationships, houses owned or rented, marriages, or job-worker matches, is a stock and the level of that stock rises and falls in response to the movements of the flows into and out of the stock. Although it is the level of a stock which is typically of interest, the analyst must focus on the determinants of the inflows and outflows.

In the flows approach, unemployment is a state of labor market participation that a worker typically occupies for a relatively short period of time while he or she seeks the complementary state of acceptable employment. Different individual workers flow into the state and others out in any period of time, and these flows vary over time. Because the outflow was roughly proportional to the level, unemployment tended toward the level that balanced the two flows. Interestingly, in the 1930s the designers of the Census of Population Survey (CPS) in the U.S., the source used to measure unemployment, already had this view. To be so classified as unemployed one must give a positive answer to the question: Have you looked for a job in the last week?

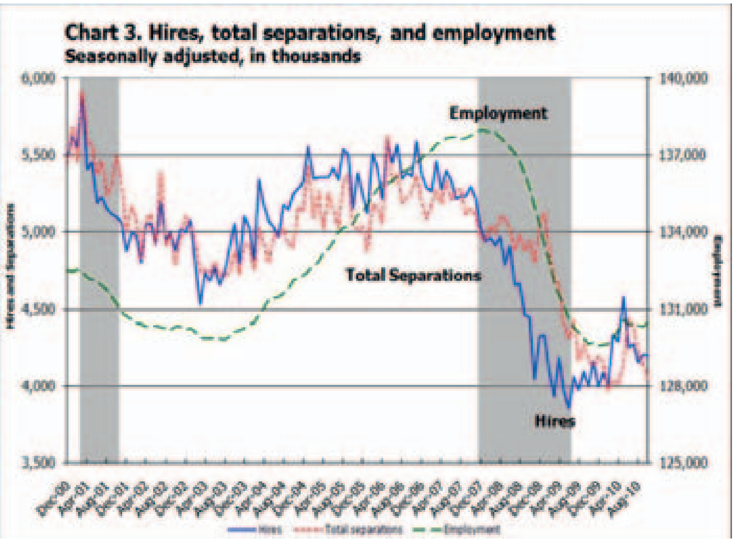

Figure 1. Hires, total separations, and employment. Seasonally adjusted in thousands. Source: Bureau of Labor Statistics, www.bls.gov.

Figure 1 provides a graphical illustration of the labor market dynamic during most of the last decade. In the figure, the flow of new hires and the flow of workers separating from employment in each month are represented by the jagged lines while the level of employment is given by the smooth curve. The height of each flow on the left vertical scale represents the rate per month in 1,000s of workers, while the employment in each month corresponds to the horizontal distance on the right vertical scale. Because the change in employment is equal to the difference between the hire and separation flows, the level of employment rises when hires exceed separations and falls when the hire flow is less than that required to replace those that leave employment. Of course, the number of unemployed moves in the opposite direction in response to the difference in the two flows. The shaded areas in the figure represent the recession periods as determined by the National Bureau of Economic Research. As the figure suggests, recessions are generally periods in which separations exceed hires. Note that the flows are always large. Although in the order of 5.5 million per month in the best months, the flow of hires is still 4 million at the depth of the Great Recession. The popular idea that there are no jobs available is simply inconsistent with the facts.[1]

Given the flow view of the labor market, finding an acceptable job is the outcome of an information gathering process in which the worker exploits contacts, friends and neighbors about job availability while employers engage in complementary recruiting activities. Other more formal channels of

1 For a more comprehensive view of the data, see Davis *et al.* (2006).

information transfer are also available. Help wanted advertising in newspapers and vacancies posted on the internet inform decisions about where to apply for specific openings. For any searching worker, the product of this effort is a potential sequence of offers distributed over time. When an employment offer arrives, the worker has to decide whether or not to accept it, given the information gathered to date and general information about the availability and generosity of offers that might be obtained in the future. In other words, the alternative to acceptance for a market participant is not home production, schooling, or some other non-market activity if the person desires a job. The default option is continued search. But the frequency of offer arrival is not a deterministic outcome subject only to a worker's choice of search effort and acceptance criterion. It is also affected by the actions of all other workers and employers in the market. How the interactions of all the participants generate allocations of workers to jobs is the issue of interest to economists.

In sum, worker-job matching takes time as a consequence of information frictions associated with the fact that no worker knows the location and terms of all jobs and no employer knows where to find every worker among those available. In this more nuanced view of the labor market, how wages are set is problematic. In this environment "competitive market clearing" in the usual sense of the term is impossible for the following reason. Because a satisfactory employment relationship justifies search and matching investment in time and money by both workers and employers, once a job opportunity is located a comparable alternative matching option is not available instantaneously to either the worker or the employer. Furthermore, because search and recruiting effort are investments and jobs have duration, it is the future streams of wages and profit associated with an employment opportunity that matter, not just their current values.

3. THE PHELPS VOLUME

In the mid 1960s a debate raged over the Phillips curve and its implication for economic policy. The curve as reported by Phillips (1958) was a negative statistical association between the inflation rate and the unemployment rate. Some interpreted the relationship as an unpleasant trade-off that policy makers had to face. According to this view, one could use policy to lower unemployment only at the expense of increasing inflation.

However, some argued that the theory behind the curve was suspect. Milton Friedman, in his AEA Presidential Address (1968), declared that the economy tends toward a 'natural' rate of unemployment determined in market equilibrium which is invariant to the inflation rate, at least in the long run. Monetary policy can reduce the unemployment rate over the relative short run at best. Any attempt to maintain the level below its natural or equilibrium rate though monetary policy will be frustrated by ever increasing inflation.

Edmund Phelps (1968a, 1968b) made a similar point. Moreover, he complemented his assertion with the new view of how markets with search friction work. Specifically, he suggested that some positive unemployment level was a natural outcome of the matching process in the labor market. Any attempt to lower unemployment below that determined by the rational agent behavior of individual employers and workers by inflating the economy would simply induce adjustment back to its natural level at a higher rate of inflation. Underlying the arguments of both Friedman and Phelps was the proposition that there is no money illusion, at least in the long run. Any attempt to affect the real wage by artificially manipulating the price level though monetary policy would induce agents in the labor market to agree to offset the effect by revising the nominal wage.

This was the environment in which I began my research and teaching career. At Northwestern we created an informal reading group in the spring of 1968 focused on the macroeconomic implications of these new ideas circulating in the profession about the labor market. I began to think about how to capture the essential features of a decentralized market with search friction in a formal but simple economic model. Using the labor market as my focus, I came up with the idea of modeling the consequence of search and matching friction as the outcome of a sequence of random meetings between potentially interested parties. After I had written a very long working paper designed to formulate and work out some of my idea, Ned Phelps asked me to present it at a conference he was organizing on the topic. I was fortunate to have this, my first major paper, "A Theory of Wage and Employment Dynamics" published in the collection of papers presented at his conference that became universally known as the "Phelps volume."

The Phelps volume, *Microeconomic Foundations of Employment and Inflation Theory*, was published in 1970. It became a classic. The basic message of the collection was that one could and should consider the dual macroeconomic problems of employment and inflation as the outcomes of individual agents' interaction, in which each behaves in his or her own interests in a market environment characterized by uncertainties and incomplete information. In other words, the authors argued that macroeconomics should be founded on microeconomic principles. In addition to myself and Phelps, the authors included a third Economics Prize winner, Robert Lucas. Although the three of us have not always agreed on the details and have taken different directions in the pursuit of the goal, we have shared a common view that macro-economics needs a foundation based on equilibrium market analysis.

My second paper, "Job Search, the Duration of Unemployment and the Phillips Curve", published in the *American Economic Review* in 1970 as well, was an attempt to use new ideas about decentralized exchange in a labor market with search friction to provide an interpretation of the empirical Phillips curve. I refer to it here, not because I accomplished that goal. In fact, I regard the paper as a failure in that dimension for several reasons. One of these was the inability to close the model with a convincing theory of agent expectations. Arguably this could have been done by invoking "rational expectations," a concept that had already been introduced by one of my Carnegie Tech professors, John Muth (1960). I did not see its relevance for my work at the time, even though I had participated in a seminar on the subject briefly as a student. The issue was subsequently resolved by later adopting the approach in spite of its drawbacks.

4. ONE-SIDED SEARCH MODELS

One formal feature of my Phelps volume paper, still incorporated into what is now regarded as the Diamond-Mortensen-Pissarides (DMP) model of unemployment, is a focus on the law of motion for the unemployment rate. The law reflects the fact that the change over time in the number of workers who are unemployed is the difference between the flow of workers who enter unemployment and the flow of those who leave. In the simple case in which workers are either employed or not, the inflow is the product of the separation rate, denoted as s, and the number of workers employed and the outflow is the product of the job finding rate, f, and the number of unemployed workers. The unemployment rate, u, tends toward the number that balances the two flows, the solution to

$$\frac{u}{1-u}=\frac{s}{f}. \quad (1)$$

Equivalently, since the inverse of the job finding rate, $1/f$, is the average duration of an unemployment spell and $1/s$ is the average duration of an employment spell, the equation asserts that the "odds of being unemployed," the left hand side of (1), equal the average length of an unemployment spell divided by the average length of an employment spell.

In my original search model of unemployment, the average length of an unemployment spell was determined by a combination of "choice and chance." In my Phelps volume paper, the worker was viewed as the recipient of a sequence of job offers that arrived at random intervals over time. The sole economic decision was whether to accept or not an available offer. Although this decision problem was discussed in my Phelps volume paper, it was fully formalized in the comparison paper, Mortensen (1970b). McCall (1970) provided similar formulations at about the same time based on a mathematical analysis of the optimal stopping problem model borrowed from stochastic decision theory.

The essential assumptions of the formal optimal stopping model as applied here are that the worker cares about the expected discounted stream of future wages offered by a job and that an offer is a random draw from the distribution of possible wage offers which is known to the worker. Given these assumptions, the decision to accept or not is analogous to the problem of exercising a stock option. Formally, the probability that the next offer is accepted is $1-F(R)$ where R denotes the worker's reservation wage and $F(w)$ is the wage offer cumulative distribution function. By definition, the reservation wage equates the expected present value of the worker's future income were he to accept the job with the value of forgoing the employment opportunity by continuing to search for an alternative offer. Given the reservation wage, the job finding rate is the product of the offer arrival rate, denoted as λ, and the probability that it is accepted, i.e., $f = \lambda(1 - F(R))$.

Once a job offer is accepted, the worker has no incentive to quit in a stationary environment. Hence, the simple theory of the unemployment level can be summarized using the stock-flow equation (1) as

$$u = \frac{s}{s + \lambda(1 - F(R))}, \quad (2)$$

where the job separation rate s is regarded as exogenous.

By implication, the unemployment rate is higher if *either* job offers arrive less frequently when the workers are unemployed *or* layoff shocks occur more frequently given the reservation wage. Of course, the reservation wage, that which equates the value of employment and unemployment for the work, should depend on these same quantities. In the case of no search-on-the-job, it is equal to the income that a worker forgoes by accepting employment, say the unemployment benefit, plus the option value of continued search,

$$R = b + \lambda \int_R^{\infty} (W(w) - U) dF(w). \qquad (3)$$

In this expression, the first term, b, represents the unemployment benefit and the second term is the expected gain in future wage income attributable to continued search, where $W(w)$ denotes the future earning associated with a job that offers wage w in the present. In words, it is the expected gain in future wages associated with the possibility that the worker will find a job offering a wage above the reservation wage were he to continue searching as an unemployed worker.

The fact that this theory had important implications for the time variation in employment induced by fluctuations in business activity and for the design of labor market policies was immediately noted. For example, it was well known that it is easier to find a job during booms and that layoffs are more frequent during business contractions. These observations could be mapped into the statement that the job finding rate f is procyclical while the employment separation rate s is countercyclical. For a given reservation wage, the source of the time variation were movements in the offer arrival rate λ and the job separation frequency s. As a consequence of these movements unemployment rose during recessions and fell during recoveries, again holding the reservation wage constant. However, equation (3) implies that the reservation rate is increasing in f and decreasing in s. That is, a rational worker should require a higher wage to forgo search if jobs are easier to find and should be less demanding if a layoff is more likely in the future. Although these movements tend to offset those of f and s, evidence suggests that the direct impacts dominate.

Virtually all economically developed and many developing countries provide unemployment insurance (UI) financed by taxes in the form of benefit payments to those who lose their jobs. The social purpose of these policies is to insure households against an adverse consumption shock associated with job loss. However, the reservation wage equation suggests that one should expect a side effect to this policy. Namely, because the UI benefit is forgone when a worker takes a job, the wage earned when employed compensates for the loss of this source of income. Hence, the reservation wage increases and the job finding rate decreases with b given the rate at which jobs are located. This disincentive has the adverse effect of raising the overall rate of unemployment. Another corollary of a richer version of the model is that a longer benefit period also decreases the job finding rate for similar reasons. There is now a vast empirical literature confirming these predictions. These findings shed light on the trade-off between the obvious social benefit of unemployment insurance and a "moral hazard" effect on the incentive to search for employment, a dilemma that is at the heart of debate over the design of unemployment compensation systems. However, it should be remembered that the same is true of any insurance program.

4.1 Equilibrium in a Two-Sided Search Model

But, the one-sided search model does not provide a complete theory of employment and wage determination useful for dynamic and policy analysis. The demand side of the market was not explicitly modeled in most papers published in the 1970s. The exception to that rule is paper on price setting by Diamond (1971).[2] In that paper, Peter asked, "How are prices set in markets with search friction when there are many buyers and sellers?" He gave a startling answer. Applying the game theoretic concept of a non-cooperative Nash equilibrium, he found that all sellers charge the buyers' reservation price. The logic behind this result is quite straightforward. First, if all suppliers offer the same price and workers know that fact, then no buyer has an incentive to search for more than one offer. If every buyer purchases at the first opportunity, sellers maximize profits by charging the reservation price. Neither worker nor employer has an incentive to deviate.

His result is now known as the "Diamond paradox" because it suggests that what appears to be a relatively small deviation from the conditions of perfect competition generates a far different conclusion about price outcomes. Later Diamond's conclusion was modified by several authors by showing that there were other equilibria, including the competitive equilibrium, if one relaxed one of his assumptions, that workers/buyers must respond to offers one at a time. Indeed, Burdett and Judd (1983) demonstrate that an even more interesting equilibrium exists that could be characterized as pure price dispersion, an outcome in which different sellers charge different prices for the same good. Later, Burdett and Mortensen (1998) point out that Burdett-Judd conditions for pure dispersion are satisfied if both employed workers and unemployed workers search sequentially without recall. Indeed, the only equilibrium in their formulation is one in which the so-called "law of one price" fails. Of course, this result simply reinforced Diamond's argument

2 It is of historical interest to note that the only reference given in Diamond's paper is to Phelps *et al.* (1970).

that imperfect information about trading opportunities led to a much richer theory of price and wage determination than that implied by the theory of perfect competition.

Research on the approach to search equilibrium pioneered by Diamond's paper is still very active. In particular, it has been developed by labor economists for the purpose of understanding wage differences across firms and individual wage growth associated with job-to-job movements observed in longitudinal data. The recent discussion of developments in the theory is contained in Mortensen (2003) and the empirical literature on the topic is reviewed by Postel-Vinay and Robin (2006). This activity is reflected in the fact that a recent issue of the *Journal of Labor Economics* was devoted to this approach to understanding monopsony in the labor market. (See Ashenfelter et al, 2010)

The model of wage setting adopted in the macro literature, however, is based on bilateral bargaining theory. In that setting, neither worker nor employer has the power to set the wage. Instead, the wage must be a mutually agreed on as the outcome of bargaining between worker and employer. The "pie" to be divided in the bargain is equal to the wedge between the marginal value of a worker to the employer and the worker's reservation wage. This wedge is positive precisely because time and resources are required to find an alternative match partner.

Naturally, the wage bargain must be struck ex post, after the two parties meet. This fact raised the following question: Would the division of match rents that resulted from bilateral bargaining provide correct incentives to participate in the search and matching process? Although the obvious answer was a resounding no, the formulation of the problem initiated a new literature on the properties of market equilibrium solutions to fully articulated search models, models that incorporated the search and matching behavior of both workers and employers.

My own early contributions in this area includes "Specific Capital and Labor Turnover" published in 1978 and two papers published in 1982: "Property Rights and Efficiency in Mating, Racing and Related Games" and "The Matching Process as a Non-cooperative Bargaining Game." These papers along with those of my co-winners Peter Diamond (1979, 1982a, 1982b) and Christopher Pissarides (1979, 1985, 1986) were among the first to formulate two-sided search models and to deal explicitly with the dual issues of the existence and efficiency of search equilibrium.

The model I used in my 1982 papers was one of "partnership formation." A partnership in this context is a coalition of two types of agents that can produce a stream of joint value that exceeded the sum of what the two partners can accomplish independently. A business partnership, a job-worker match, and a marriage are among the examples of partnership. The structure of the problem of partnership formation is closely related to the Becker (1973) marriage model with transferable utility extended to include search and matching friction. As in existing models of "one-sided" search, potential partners are assumed to meet at random. Although all partnerships could be viewed as indistinguishable, "match quality," the output specific to the match, is a random variable revealed only after two agents meet. This feature of the model distinguished it from the case in which agents have fixed "abilities" or other characteristics that determined joint output.

As already noted, search and matching friction implies the existence of a surplus once a compatible pair meets. In all papers referred to above, the division of the surplus is jointly rational in the sense that no other exists that makes both parties better off by assumption. Hence, matches form if and only if in the interest of both parties. Although there are generally many jointly efficient divisions of the surplus as in any bilateral bargaining problem, it was assumed that one was chosen by every pair that could be characterized by the worker's share of the rent, a number which some interpreted as "worker bargaining power."

The concept of a matching function is an explicit market level relationship that purports to characterize the outcome of the process by which agents meet and match. Along with a bargaining approach to wage determination, Diamond and Maskin (1979) also used the concept of a matching function in their earlier paper entitled "An Equilibrium Analysis of Search and Breach of Contract I" published in the *Bell Journal of Economics*. Their paper had extended the analysis in my own *Bell Journal* (1978) article by setting the latter in a market equilibrium context.

The appearance of our papers on two-sided search led to communication between Peter and me regarding the development of our ideas over the next several years. I think it is fair to say that we fed off each other in a very collaborative fashion during the early 80s even though we did not coauthor. Indeed, Peter's valuable commentary on it was published along with my paper "The Matching Process as a Non-cooperative Bargaining Game" in the same collection. Also in the same year, Peter published his own classic contributions to two-sided search in "Wage Determination and Efficiency in Search Equilibrium" and "Aggregate Demand Management in Search Equilibrium." Pissarides had already broken new ground in studying the macroeconomic implications of the flows approach to labor market analysis by using the matching function as a tool to study equilibrium unemployment in his article "Job Matching with State Employment Agencies and Random Search," published in 1979.

The matching function relates the flow of worker-job meetings to the number of workers unemployed, the number of job vacancies, and the intensities with which workers search and employers recruit. It successfully captures the key implications of search frictions that prevent an instantaneous encounter of trading partners. It proved a particularly powerful tool for modeling two-sided search frictions that stem from information imperfections about potential trading partners, not least because it can be incorporated into models of unemployment without adding excessive complexity.

In his article "Short-Run Equilibrium Dynamics of Unemployment, Vacancies, and Real Wages" published in 1985, Pissarides used the matching function to highlight the effects of cyclical fluctuations in demand on unemployment and wage dynamics. Although I was aware of and followed his work with interest in these years, only after the publication of the first edition of his book *Equilibrium Unemployment* in 1990, which fully articulated the first generation of the DMP model, did we collaborate in a string of coauthored papers, initiated by "Job Creation and Job Destruction in the Theory of Unemployment," published in 1994. This paper extended the model in the first edition of his book to include endogenous job separation as well as creation. Subsequently, the extended model was incorporated in the second edition of the book, published in 2000.

In my own two-sided search models, I considered two different specific but plausible functional forms for the matching function: linear and quadratic. Both can be generated as descriptions of simple sampling schemes sometimes referred to as "ball and urn" models. One can think of the linear matching function case as one in which each unmatched agent on either side of the market can randomly contact an unmatched agent on the other side at a chosen and costly frequency. Of course, the same agent might also be contacted by someone of the complementary type in the market. Hence, the aggregate instantaneous rate at which jobs and workers meet is a linear function of the numbers of searching agents on the two sides of the market. By implication, the expected frequency of meetings for a worker, which is the ratio of the aggregate rate divided by the number of unemployed workers given random search, can be viewed as a Poisson random variable.

A "quadratic" matching function is an extension of this simple set-up, in which the aggregate meeting rate is proportional to the product of the numbers of the searching agents of the two types. Indeed, this case can be viewed as one in which each individual contacts another on the other side but does not know whether the potential partner is matched or not. Hence, the effective meeting rate with an unmatched agent of the opposite type is the product of the numbers of the two agent types. In the labor market context, the linear case can be represented as $M(u, v) = a_1 v + a_2 u$ where u and v denote the numbers of unemployed workers and vacant jobs respectively and a_1 and a_2 are constants reflecting the frequency of contacts. Similarly, $M(u, v) = (a_1 + a_2) uv$ in the quadratic case.

In my two 1982 papers, an equilibrium solution is a search strategy for every participant that maximizes the agent's expected future income given the strategies of all the other participants. Formally, these are the conditions required of a Nash equilibrium to a non-cooperative game of two-sided search, a concept developed by John Nash, made famous to the general public in the movie "A Beautiful Mind." I was able to show that a unique Nash equilibrium exists in the linear matching function case and at least one symmetric equilibrium exists in the quadratic matching case as well.

A solution to the model, a choice of a search strategy by every participant, is constrained efficient if there is no other feasible set of strategies that yields a higher level of aggregate income given the frictions in the model. As no agent in the economy owns the matching process, the question of economic efficiency of a two- sided search equilibrium solution arose rather naturally.

Who should receive the ex-post quasi-rent associated with an acceptable match, given that both parties invest in its formation?

There are two so-called external effects present in the model that are neglected in the decision calculus of any agent. More recruiting effort by employers benefits workers by reducing the time it takes to find a job, while more unemployed worker search increases it. However, another vacant job in the market tends to increase the time it takes to match all the other vacancies. When deciding whether or not to create a new vacant job, employers do not take account of these costs and benefits. Hence, there is a reason to believe that an equilibrium will not be efficient.

After demonstrating the existence of a unique equilibrium for any match surplus sharing rule in the linear case, I proved that there was one that would induce efficient incentives. But the result did not hold in the quadratic case. Later Hosios (1990) generalized my results by demonstrating that a unique value of the worker's share of match surplus exists that maximizes equilibrium aggregate incomes if and only if the matching function $M(u, v)$ is increasing, concave, and homogeneous of degree one in its arguments. Interestingly, the efficient worker's share is the elasticity of the matching function with respect to the number of unemployed workers. Subsequently, Moen (1997) showed that this condition always held in a variant of the model in which workers can observe the wages offered by individual employers but are uncertain about whether or not a particular job is open.

The possibilities are more interesting in the quadratic case. First, it follows as a corollary of the Hosios theorem referred to above that there is no way to share match surplus so as to provide both workers and employers with incentives that induce an efficient equilibrium outcome in this case. Indeed, one can construct examples in which there is "too much" and others in which there is "too little" unemployment. Furthermore, Diamond (1982a) demonstrated that there can be multiple equilibria that are ranked according to aggregate welfare. That is, different equilibria can exist each supported by different rational expectations about the future, and these are Pareto ranked. In one, employers are optimistic about the future and as a consequence there are many jobs created, unemployment is low and economic welfare is high. While in another equilibrium, employers are pessimistic, vacancy creation is low, aggregate income is low and unemployment is high. Later Diamond and Fudenberg (1989) elaborated the dynamics of the original model when the multiplicity occurs because of increasing returns to scale in the matching process, and in Mortensen (1999) I provided a complementary analysis for the case of multiplicity generated by increasing returns in production.

Some economists view the coexistence of different possible outcomes as a problem with the theory. However, game theory has taught us that multiple equilibria naturally arise in games of coordination. Could our lack of employment recovery from the current recession imply that the economy has found a "bad" outcome as a consequence of pessimistic expectations? It is a theoretical possibility.

The specification of the matching function adopted in what became known as the DMP model is a generalization of the linear case defined above. Specifically, the flow of matches is represented increasing, concave, and linearly homogenous in its two arguments, the number of unemployed workers and the number of vacancies participating in the matching process. Later empirical studies provided support for these assumptions (Petrongola and Pissarides, 2001). Given the specification, the expected meeting rates per unemployed worker can be expressed as

$$\lambda = \frac{M(u,\upsilon)}{u} = M(1,\theta) \equiv m(\theta)$$

where $m(\theta)$ is an increasing and concave function of the ratio of vacancies to unemployed workers $\theta = u/u$, known in the literature as "labor market tightness." Hence, unemployment should tend over time toward the steady state value

$$u = \frac{s}{s + m(\theta)} \qquad (4)$$

since the job finding rate in equation (1) is the meeting rate and there are no shocks to the environment. In other words, the steady state unemployment rate falls as the number of vacant jobs increases relative to the number of unemployed workers searching for these jobs.

Chris Pissarides (1985) pointed out that this relationship "explains" the empirical Beveridge (1944) curve, the negative association between unemployment and vacancies which seems to hold in all market economies. Recently, Shimer (2005) demonstrates that the critical assumption required, that the job finding rate is increasing in the ratio of vacancies to unemployment, holds clearly in the data. Indeed, after the data are appropriately filtered to wash out trend, Shimer finds that the log of the job finding rate is closely approximated by the log of the vacancy-unemployment ratio with a positive slope while the log of the vacancies is a roughly linear in the log of the unemployment rate with a negative slope. Both the Beveridge curve and the relationship between the job finding rate and the tightness for the post WWII period in the U.S. are illustrated in Figure 2.

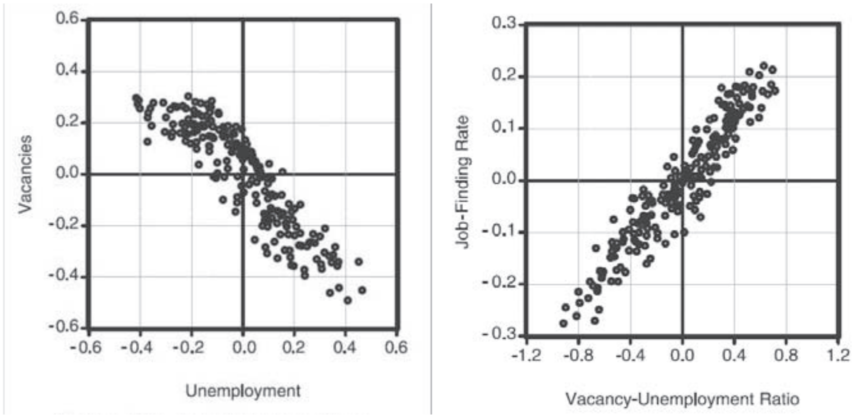

Figure 2. Quarterly US Beveridge Curve, 1951–2003 (left) Monthly US Matching Function, 1951–2003 (right). Source: Shimer (2005).

Wage determination in the DMP model is based on bilateral bargaining theory. Specifically, the worker receives some share $0 < \beta < 1$ of the match surplus, the difference between the expected present value of future match output and the value of unemployed search, and the employer receives the complementary share $1 - \beta$ as future profit. The particular share β was left unspecified. Its value reflects worker "bargaining power" but is not explicitly modeled in the theory. Formally, the expected present value of future wages, denoted as *W*, is given by

$$W = R + \beta\, S(R)$$

where *R* represents the value of unemployment, the worker's reservation value, and $S(R)$ is the surplus value of a actual job match expressed as a function of the worker's reservation wage *R*.

Since vacancies reflect employer demand for labor services, the rational for the Beveridge curve relationship makes sense once search and matching friction is taken into account. However, a clear and simple statement of labor demand consistent with the search and matching framework was not fully articulated until Pissarides (1985) in his *American Economic Review* article introduced the "free entry" condition. Given that it takes time to fill a vacancy, the appropriate demand condition requires that the expected cost of posting a vacant job, the product of the flow cost of holding a vacancy open, represented as k, and the expected time required to fill it, which is $\theta / m(\theta)$, must be equal to the present value of the future profit that employers can expect from a job-worker match. Hence, the free entry condition is

$$\frac{k\theta}{m(\theta)} = (1-\beta)S(R) \qquad (5)$$

in a simple stationary environment where the left side is the expect cost of filling a vacant job and the right side is the firm's share of the surplus value of a job.

For any surplus value of a match, the model is complete. However, a precise specification of $S(R)$ is problematic because it exists only in the minds of the employer and employee involved in the match, and even they need not agree. Formally, if they do agree on how to forecast the future, it is the current or present value of the expected flow of future income to both worker and employer, were they to continue their match, less the expected present value of income given they were to search instead. Because every employer is indifferent between search and continuation when the free entry condition is satisfied, $S(R)$ is the present value of the expected future stream of differences between the market value of their joint product flow and the worker's value of search. Therefore, one can infer that the function $S(R)$ is decreasing in the reservation wage R.

Otherwise, exactly what the future will bring in terms of both the market for the product of a particular match as well as the appropriate rates at which to discount the future surplus is fraught with uncertainty. How do participants in the labor market form these expectations? Is it possible to read the minds of human agents in our models?

The problems of expectation formation concerning future economic prospects were clearly understood by Keynes (1936). In his own analysis, which included two full chapters in *The General Theory of Employment, Interest and Money,* he emphasized the role of "business confidence" and "animal spirits". However, modern macroeconomics has side-stepped the issues that he thought relevant by adopting the hypothesis of "rational expectation." The original proponent of this approach, John Muth (1960), conjectured that agents in any market know how prices are set and use this knowledge to form expectations. Although this statement sounds reasonable in theory, in practice the agents must have a model in mind and that model must be the one they populate. Although this is the only way in which to create a closed consistent model of market behavior, the postulate is ridiculously self-referential. Nonetheless, I along with others have adopted it in our papers.

Recent events has called into question the hypothesis of "rational expectations" and a related concept, that of "efficient markets." I too have become far more skeptical. However, I wish to argue that the hypothesis is not the principal contribution of the DMP model to our understanding of how labor market work. For these reasons, I leave $S(R)$ unspecified. I made only two assumptions: First that worker and firm have the same expectations about future prospects. Second, one can represent these common expectation as a single number, the surplus value of a match, $S(R)$, which is decreasing in the reservation wage. Then, the entry conditions and the reservation wage equation combine to yield,

$$R = b + m(\theta)\beta S(R) = b + \frac{\beta k \theta}{1-\beta}. \quad (6)$$

In short, the reservation wage is linear and increasing in market tightness with an intercept determined by the unemployment benefit and a slope coefficient that reflects the cost of posting a vacancies and the relative share of future returns that flow to the worker. The reservation wage rises with market tightness because an increase in the vacancy-unemployment ratio decreases the time it takes to find an alternative job.

A search equilibrium in the DMP model can now be represented is a reservation wage R and a vacancy-unemployment ratio θ that satisfies the

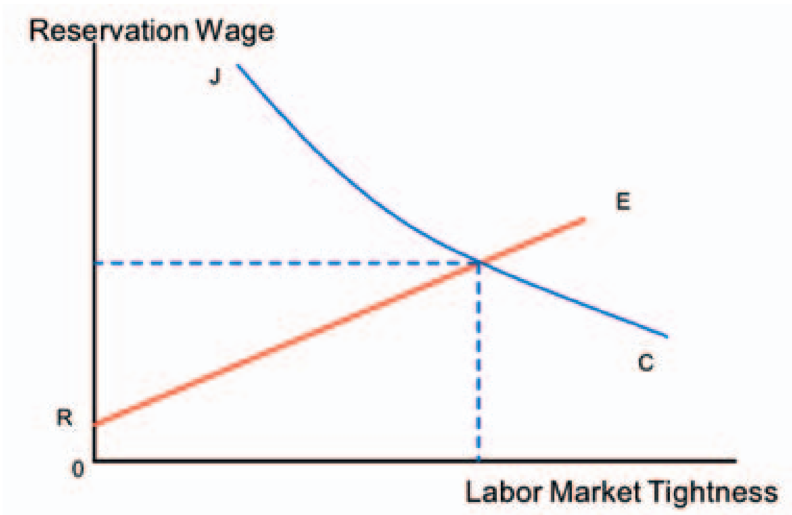

Figure 3. Search Equilibrium.

free entry or job creation condition (5) and the reservation wage equation (6). Since the former defines a downwards sloping relationship between θ and R, which is represented by the curve JC for "job creation" in Figure 3, while the wage equation, RE in the Figure, is upward sloping there can be only one solution at the intersection of the two curves. Furthermore, both R and θ, are positive if and only if match product exceeds the unemployment benefit, $p > b$.

Now, what does all this algebra and geometry imply? To bring this question closer to home, how does the framework that we three created help the rest of us to understand the current unemployment crisis in the U.S. and many countries in Europe? Well, suppose for some reason employers' expectations about future profitability tanks. Let us say that this negative "shock" to employer-employee expectations occurs because households suddenly and sharply reduce expenditures on goods and services in response to a decrease in the prices of their homes. In response to the drop in wealth suffered as a consequence of the 2008 financial crisis, homeowners and firms did attempt to increase savings in financial assets by reducing expenditure on durables. Expecting persistence in this process, worker and employers revised downward their expectations about the present surplus value of a future match, particularly that which was designed to produce those durables. Hence, the model predicts a sharp and large fall in the value of the right hand side of the free entry condition represented in equation (5). Employers respond by creating fewer vacancies. As a result, the vacancy-unemployment ratio required by the equality of the left and right sides of the free entry condition falls. In geometric terms, the initial impact corresponds to a downward shift of the JC curve as illustrated in the Figure 4.

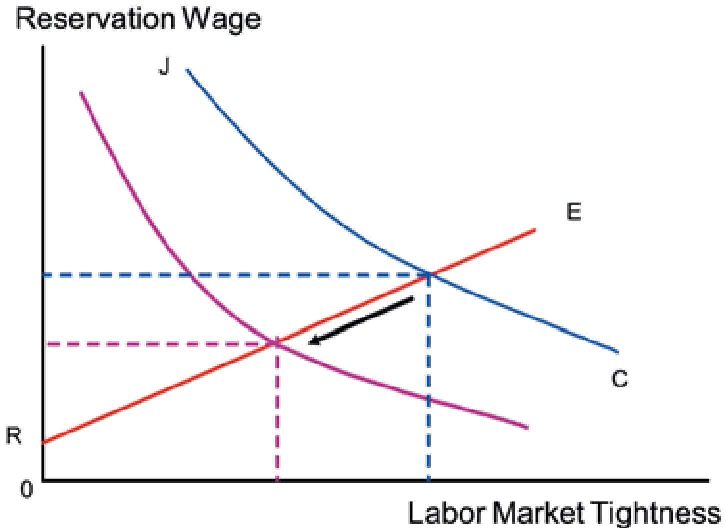

Figure 4. Negative Expectation Shock.

Another factor contributing to the downshift in the JC curve involves the reaction of commercial banks to the financial crisis. Initially they suspended the lines of credit of smaller businesses customers used by them to finance payroll and refused to extend similar loans to new customers. All of a sudden, their willingness to bear risk vanished. Later when they did resume lending, the banks applied more stringent standards for the loans that were granted and charged higher interest rates. Although it is true the Federal Reserve was able to reduce the nominal rates that federal banks paid for reserves, the spread between that rate and the effective interest rates charged lenders has increased and remained large, as Robert Hall documented in his 2010 AEA Presidential address. This rise in the cost of the capital required to create jobs contributed to the drop in the present value of a new job-worker match. As the spread has not diminished, it is also one of the reasons for a slow recovery of vacancies since the end of the recession.

The reduction in market tightness does have a negative effect on wages of new hires, to the extent that the joint value of a match expected by workers and employers is shared. This effect corresponds to a movement along the RE curve in Figure 4 induced by the downward shift the JC curve. The net result is that the lower wage will cushion but not offset the impact of the negative expectation shock on tightness. Finally, because tightness falls, the unemployment rate must rise given the linkage represented by Beveridge curve relationship represented in equation (4).

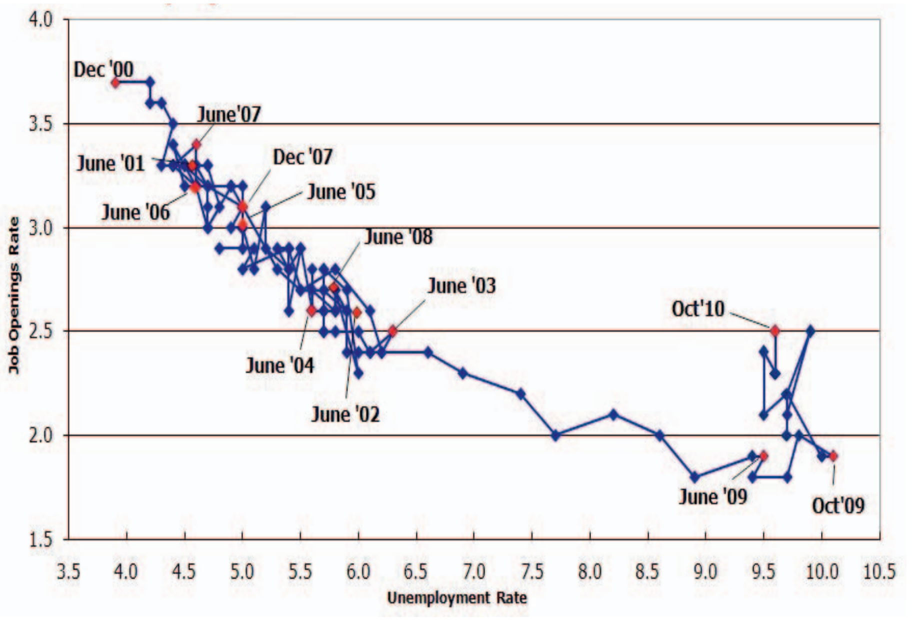

Figure 5. The Beveridge Curve (job openings vs. unemployment rate) Seasonally adjusted. Source: Bureau of Labor Statistics, CES and JOLTS data, www.bls.gov.

Are these predictions consistent with what has happened? Figure 5, which represents the empirical Beverage curve in the U.S. since December 2009, clearly illustrate that they are. In that picture, the vacancy-unemployment combinations in the upper left corner of the diagram correspond to monthly observations from December 2000 to the beginning of the Great Recession in December 2007. Their variation reflects first the downturn movements and then the recovery from the recession of 2001. The remaining points are for the months up to and including October 2010. As one can see, the U.S. economy marched right down the Beveridge curve, roughly until the National Bureau of Economic Research declared the end of the recession in June 2009.

For the last year, the unemployment rate has hovered about 9.5 to 10 percent of the work force even though vacancies have increased somewhat, but only to levels associated with the trough of the previous recession. Some wonder whether these facts reflect the development of "structural unemployment," a mismatch between the jobs that are available and the qualifications of the workers seeking them. It is true that jobs in construction and manufacturing were the first to be destroyed in the downturn, but these interest rate-sensitive industries are always hardest hit in a recession. In other words, the great recession in no different in this respect. The fact that it followed a major financial crisis is different. As Robert Hall (2011) argues in his recent AEA Presidential address, the attempt by the Federal Reserve to lower interest rates by reducing the rate at which federal banks can borrow reserves has not had the intended effect on either the lending of commercial banks or the interest rates they charge to employers and consumers. Banks are still reluctant to make loans, and the spread between the rate at which banks borrow and lend has widened. As the cost of capital is a critical determinant of the expected value of a match, the incentive to create new jobs has not recovered to the extent needed to lower unemployment to anything like a normal level.

A recovery will happen only when employers believe that hiring workers now will be profitable. Government policy will assist or harm the chances of recovery to the extent it impacts the expectations about these prospects. The following are corollaries: (1) Fiscal tightening is doing harm because it reduces the expected demand for goods and services supplied now. (2) It is essential to find a way to get commercial banks to return to their role of supplying the working capital needed by small business to finance job creation at reasonable rates of interest. (3) Now, while the resources are available, is an excellent time for the government to invest in future productivity and growth.

REFERENCES

1. Ashenfelter, O.C., H. Farber, and M.R. Ransom (2010), "Labor Market Monopsony," *Journal of Labor Economics*, 28: 203–210.
2. Becker, G. (1973), "A Theory of Marriage, Part 1," *Journal of Political Economy*, 81: 813–846.
3. Beveridge, W.H. (1944), *Full Employment in a Free Society*, George Allen and Unwin.
4. Bontemps C., J.-M. Robin, and G.J. van den Berg (1999), "An Empirical Equilibrium Job Search Model with Search on the Job and Heterogeneous Workers and Firms," *International Economic Review*, 40: 10391074.
5. Burdett, K. (1978), "Employee Search and Quits," *American Economics Review*, 68: 212–220.
6. Burdett, K. and K. Judd (1983), "Equilibrium Price Dispersion," *Econometrica*, 51: 955–970.
7. Davis, S.J., J. Haltiwanger, and S. Schuh (1998), *Job Creation and Destruction*, MIT Press.
8. Davis, S.J., R.J. Faberman, and J. Haltiwanger (2006), "The Flow Approach to Labor Markets: New Data Sources and Micro–Macro Links," *Journal of Economic Perspectives*, 20: 3–26.
9. Diamond, P.A. (1982a), "Aggregate Demand Management in Search Equilibrium," *Journal of Political Economy*, 90: 881–894.
10. Diamond, P.A. (1982b), "Wage Determination and Efficiency in Search Equilibrium," *Review of Economic Studies*, 49: 217–227.
11. Diamond, P.A. and D. Fudenberg (1989), "Rational Expectations Business Cycles in Search Equilibrium." *Journal of Political Economy*, 97: 606–619.
12. Diamond, P.A. and E. Maskin (1979), "An Equilibrium Analysis of Search and Breach of Contract I: Steady States," *Bell Journal of Economics*, 10: 282–213.
13. Friedman, M. (1968), "The Role of Monetary Policy," *American Economic Review*, 58: 1–17.
14. Hall, R. (2011), "The Long Slump," forthcoming in *American Economic Review*. Available online at http://www.stanford.edu/~rehall/.
15. Holt, C.C. and M.H. David (1966), "The Concept of Vacancies in a Dynamic Theory of the Labor Market," *Measurement and Interpretation of Job Vacancies*, 73–141. National Bureau of Economic Research, New York.
16. Hosios, A.J. (1990), "On the Efficiency of Matching and Related Models of Search and Unemployment," *Review of Economic Studies*, 57: 279–298.
17. McCall, J.J. (1970), "The Economics of Information and Job Search," Quarterly Journal of Economics, 84: 113–126.
18. Moen, E. (1997), "Competitive Search Equilibrium," *Journal of Political Economy*, 105: 385–411.
19. Mortensen, D.T. (1970a), "A Theory of Wage and Employment Dynamics" in E.S. Phelps *et al.*, *Microeconomic Foundations of Employment and Inflation Theory*, W.W.Norton & Company.
20. Mortensen, D.T. (1999), "Equilibrium Unemployment Dynamics," *International Economic Review*, 40: 889–914.
21. Mortensen, D.T. (1970b), "Job Search, the Duration of Unemployment and the Phillips Curve," *American Economic Review*, 60: 847–862.
22. Mortensen, D.T. (1978), "Specific Capital and Labor Turnover," *The Bell Journal of Economics*, 9: 572–586.
23. Mortensen, D.T. (1982a), "Property Rights and Efficiency in Mating, Racing and Related Games," *American Economic Review*, 72: 968–979.
24. Mortensen, D.T. (1982b), "The Matching Process as a Noncooperative/bargaining Game," in J.J. Mccall, ed., *The Economics of Information and Uncertainty*, 233–254, NBER Conference Volume.
25. Mortensen, D.T. and C. A. Pissarides (1994), "Job Creation and Job Destruction in the Theory of Unemployment," *Review of Economic Studies*, 61: 397–415.

26. Muth, J. (1961), "Rational Expectations and the Theory of Price Movements," *Econometrica*, 29: 315–335.
27. Petrongola, B. and C.A. Pissarides (2001), "Looking Into the Black Box: A Survey of the Matching Function," *Journal of Economic Literature*, 38: 390–431.
28. Phelps, E.S. et al. (1970), *Microeconomic Foundations of Employment and Inflation Theory*, W.W.Norton & Company.
29. Phelps, E.S. (1968), "Money-wage Dynamics and Labor-market Equilibrium," *Journal of Political Economy*, 76: 254–281.
30. Phillips, A.W. (1958), "The Relationship between Unemployment and the Rate of Change in Money Wage Rates in the United Kingdon, 1862–1957," *Economica*, 25: 283–299.
31. Pissarides, C.A. (1979), "Job Matchings with State Employment Agencies and Random Search," *Economic Journal*, 89: 818–833.
32. Pissarides, C.A. (1985), "Short-Run Equilibrium Dynamics of Unemployment, Vacancies and Real Wages," *American Economic Review*, 75: 676–690.
33. Pissarides, C.A. (1986), "Unemployment and Vacancies in Britain," *Economic Policy*, 3: 499–540.
34. Pissarides, C.A. (1990), *Equilibrium Unemployment*, 1st edition, MIT Press.
35. Pissarides, C.A. (2000), *Equilibrium Unemployment*, 2nd edition , MIT Press.
36. Postel-Vinay, F. and J.M. Robin (2006), "Microeconometric Search-Matching Models and Matched Employer-Employee Data," in Blundell R., Newey W., and Persson T., eds., *The Proceedings of the 9th World Congress of the Econometric Society*, Cambridge University Press.
37. Shimer, R. (2005), "The Cyclical Behavior of Unemployment and Vacancies," *American Economic Review*, 95: 25–49.
38. Stigler, G.J. (1961), "The Economics of Information," *Journal of Political Economy*, 69: 2134–225.
39. Stigler, G.J. (1962), "Information in the Labor Market," *Journal of Political Economy*, 70: 94–105.

Portrait photo of Professor Mortensen by photographer Ulla Montan.