

14 MODELS OF UNEMPLOYMENT

Chapter 14 Contents

A. Topics and Tools	2
B. Defining Unemployment	4
<i>The statistical definition.....</i>	<i>4</i>
<i>Problems with the statistical measures</i>	<i>5</i>
<i>Natural and cyclical unemployment.....</i>	<i>6</i>
C. Introduction to Theories of Unemployment.....	8
D. Minimum Wages and Unemployment	10
<i>A simple minimum-wage model</i>	<i>10</i>
<i>Minimum-wage effects on skilled and unskilled labor.....</i>	<i>12</i>
E. Unemployment Insurance and Length of Job Search.....	14
<i>A simple model of job search.....</i>	<i>14</i>
<i>Unemployment benefits and search duration.....</i>	<i>15</i>
<i>Optimal search duration</i>	<i>17</i>
F. Unions and Unemployment	18
<i>An economy-wide labor union</i>	<i>18</i>
<i>A two-sector model of unions and unemployment</i>	<i>19</i>
G. Sectoral Shifts and Unemployment.....	20
H. Understanding Romer's Chapter 10.....	22
<i>Efficiency-wage models</i>	<i>22</i>
<i>The Shapiro-Stiglitz model.....</i>	<i>25</i>
<i>Implicit-contract models.....</i>	<i>30</i>
<i>Insider-outsider models</i>	<i>31</i>
<i>Search and matching models</i>	<i>31</i>
I. Suggestions for Further Reading.....	32
<i>On European unemployment.....</i>	<i>32</i>
<i>Theoretical models.....</i>	<i>32</i>
J. Works Cited in Text	33

A. Topics and Tools

Unemployment is one of the most visible indicators of economic activity. The rate of unemployment typically rises considerably during recessions then falls as the economic recovers. People commonly view the typical unemployed worker as suffering long-lasting despair and destitution, so the media publicize high unemployment as a great social problem. We shall see that this view of the unemployed worker is not an accurate depiction of the vast majority of those out of work in the United States. In contrast, most of the unemployed find work relatively quickly. While their income loss is significant, it is not catastrophic for most workers who suffer an unemployment spell.

Some degree of unemployment is socially and perhaps personally desirable. Much of unemployment in the United States consists of new entrants to the labor market seeking their first job, individuals who are voluntarily changing jobs or occupations, and people in jobs for which periodic or seasonal layoffs are normal, expected, and compensated for by higher wages during periods of employment. For these individuals, unemployment is not a problem at all. It is merely part of the natural functioning of a flexible and efficient labor market.

Economists often view unemployment as one facet of an inevitable process of search in the labor market. Jobs and workers are heterogeneous along many dimensions. Workers differ (among other ways) by intelligence, creativity, education, training, experience, physical size and strength, manual dexterity, ability to sustain repetitive tasks, and preferences about their work environment. Jobs vary in the abilities, education, and experience that are required to perform them, as well as in working conditions, location, opportunities for advancement, and many other characteristics. Since workers and jobs are so heterogeneous, the process of matching the characteristics of a particular unemployed worker with the most suitable vacant job often cannot be accomplished quickly. Instead unemployed workers and employers having vacant jobs engage in a two-sided search, seeking to achieve a good match as quickly as possible. The length of this search process for a typical unemployed worker is a major factor in determining the unemployment rate.

One can imagine an economy in which this matching problem could be solved trivially. If all workers and jobs were identical, for example, there would be no gain to searching for a better match. Or if everyone had instantaneous and perfect information about the characteristics of all workers and jobs, searches could be accomplished in just a moment. However, in an economy in which the matching problem cannot be solved trivially it is generally desirable to have both a positive unemployment rate and a positive job vacancy rate. Successful matching requires a pool of searching workers on one side of the market and a pool of available jobs on the oth-

er. The socially optimal unemployment rate depends on the size of the pool that is required in order for optimal matching to occur.¹ The optimal pool size, in turn, depends on the efficiency of the “matching technology” in the economy as well as on a variety of social and policy variables.

If the costs and benefits of search are largely internal to the workers and firms doing the searching, we might expect that a competitive market economy would gravitate toward the socially optimal amount of search. However, the labor markets of modern economies contain many distortions that might create externalities in the search process, causing the long-run equilibrium unemployment rate (the so-called natural rate of unemployment) to be higher or lower than the optimal rate. Imperfect access to credit markets for unemployed workers may raise the private cost of search above the social cost, shortening searches and potentially lowering natural unemployment below the optimal rate. Also impacting search are a wide variety of government policies that affect the search decisions of workers and employers, including unemployment-insurance programs, job-protection legislation, and “active” labor-market policies such as job-placement assistance and training.

Recent analysis of unemployment has focused intensely on one particular empirical problem: extremely high unemployment in continental Europe since 1980. From 1950 until 1970, the unemployment rate in most European countries averaged about 2%, roughly half of the rate in the United States during that period. Since 1980, much of Europe has suffered unemployment rates consistently in the 8 to 12 percent range, about twice the U.S. rate. Moreover, long-term unemployment (of more than one year) has been far more common in Europe than in America, although the recent recession has caused a significant increase in long-term unemployment in the United States. Theories of unemployment are able to explain some, but not all, aspects of the divergence of unemployment behavior between Europe and America.

The most interesting new analytical tool that we employ in our analysis of unemployment is dynamic programming, which is a common method of analysis of models involving transition in continuous time between alternative discrete states. In the case of unemployment models, the main states are employed and unemployed. Dynamic programming has a complicated side and a simple intuition. We’ll focus on the latter here.

¹As we shall see below, optimal matching does not generally mean making the best matches that are conceivably possible. Better matches involve costs (longer unemployment spells) as well as benefits (better fit between jobs and workers). Optimal matching balances these marginal costs and benefits.

B. Defining Unemployment

The statistical definition

In order to measure unemployment, economists have adopted a statistical definition that is only partially understood by the general public. Unemployment statistics in the United States are compiled by the Bureau of Labor Statistics (BLS) from the monthly Current Population Survey (CPS) conducted by the Bureau of the Census. This survey of approximately 64,000 households asks general questions about the labor-market status of adult members of the household during the week prior to the week in which the survey is taken.²

Based on the responses to the CPS questions, every adult is placed into one of three categories: *employed*, *unemployed*, or *not in the labor force*. Anyone who worked for pay at all during the reference week is considered to be employed, including part-time workers. Among those not employed, those who were both actively seeking work and immediately available for work, plus those who were awaiting recall from a temporary layoff from their previous job, are classified as unemployed. Anyone else, i.e., those who did not work, were not on layoff, and either were not actively seeking work or were not available for work, are considered to be out of the labor force.

In the United States, the extent of unemployment is commonly expressed as the *unemployment rate*, which is the number unemployed divided by the total labor force, which consists of the sum of employment and unemployment. (In Europe, the “headline number” is more likely to be the number of unemployed rather than the rate.) The unemployment rate ignores completely those who are classified as out of the labor force—they enter neither the numerator nor the denominator.

To the extent that it is difficult to distinguish between people who are unemployed and those out of the labor force, this may cause ambiguity in the meaning of the unemployment rate. To account for this potential problem, economists sometimes prefer the *employment/population ratio* to the unemployment rate. This ratio measures the share of the adult population that is employed and treats unemployment and out of the labor force as equivalent states.

²There is an obvious benefit to basing unemployment statistics on a survey such as the CPS rather than using such measures as applications for unemployment insurance benefits. If eligibility for benefits depends on their answer, then those out of work have a strong incentive to lie about whether they are actively seeking work. The CPS approach is likely to elicit more honest answers from respondents.

Problems with the statistical measures

Apart from the difficulties associated with all surveys, such as inaccuracy in responses, the categorization of the population by labor-market status poses some particular difficulties. There are some “gray areas” between the three categories that have led some economists (and politicians) to question the relevance of published measures.

It may seem like the least controversial boundary would be that between employment and the other categories. Individuals are either working or not, so this classification seems easy. However, if there are individuals who are working part-time because they have not been able to find a full-time job, then some degree of “problem underemployment” is masked in the statistics. Part-time work accounts for a large and growing share of employment in some OECD countries. Among countries of the European Union, over 15 percent of workers in the Netherlands, Belgium, the United Kingdom, Germany, Denmark, and Spain work part-time, while less than 10 percent of those in Greece, Portugal, Italy, Ireland, and France do.³ In many cases, laws that treat part-time workers differently than full-time workers help to explain these differences in rates.

In a study of the working-hour preferences of EU workers in 1989, a large majority in most countries indicated satisfaction with their working status.⁴ Among those working full-time, 21 percent indicated that they would rather be working part-time while 77 percent were happier with full-time work. (The remainder responded “no reply.”) Among part-time workers, 66 percent were content with part-time work while 30 percent would have preferred full-time employment. Thus, it appears that for European countries about one-third of part-time employment is involuntary, while about one-fifth of full-time workers would prefer to work part-time.⁵

This result suggests that defining employment to include both part-time and full-time workers masks problems in both directions: workers working more hours than they wish and workers working fewer hours than they want. Since there are somewhat more workers who work less than their desired amount, there is probably on net some degree of hidden underemployment in part-time employees. However, this may be less than one might have expected from the raucous reactions of union leaders and some politicians to the rise of part-time employment in recent decades.

³A detailed analysis, from which these statistics are drawn, can be found in Chapter 6 of OECD (1994). In particular, see Table 6.14 on page 93 of Volume II.

⁴This study is cited in OECD (1994), Table 6.14.

⁵These averages mask considerable variation across countries. In Greece, Spain, and France, well over half of part-time workers preferred full-time (78 percent in Greece!), while less than 10 percent of part-timers in the U.K., Denmark, and Germany wanted full-time status.

Another difficult boundary is that between unemployed and out of the labor force. Some countries have begun to collect data on *discouraged workers*, who are officially classified as out of the labor force. Discouraged workers have given up job search because they do not believe they can find a job. They are clearly part of the unemployment problem—if *their assumption is correct*. However, because they are not actively seeking work, it is impossible to tell whether they would have, in fact, been able to find a job had they continued their searches. As of 1991, the number of discouraged workers as a percentage of the labor force ranged from virtually nil in Spain, Denmark, France, and Portugal to over 3 percent in Italy, Finland, Belgium, and Japan. The U.S. number was around one percent.⁶

The presence of discouraged workers suggests that the measured unemployment rate may understate the true magnitude of unemployment. However, there may be a counterbalancing effect due to “low-intensity searchers.” Anyone without a job who answers affirmatively to the question “Are you looking for work?” is classified as unemployed. Some individuals may answer yes to this question even if their job search consists of sitting at home and waiting for the phone to ring. (Or, in some countries, working at an unreported, black-market job.) If some of the people counted as unemployed are actually, in effect, either employed or out of the labor force, then unemployment may be overstated.

Although these problems may compromise the accuracy of unemployment statistics, it is unlikely that the amount and direction of bias in the statistics change systematically from month to month, which means that changes in the unemployment rate are likely to be fairly accurate representations of changes in labor-market conditions. Moreover, information about discouraged workers and involuntary part-time workers are collected periodically to allow analysts to assess whether biases are changing over time. Interpreted with caution, survey-based measures of unemployment are useful in measuring labor-market conditions over time.

Natural and cyclical unemployment

We often make a theoretical distinction among several categories of unemployment, although it is difficult or impossible to decompose our empirical measure in a corresponding way. The most fundamental distinction is between *natural unemployment* and *cyclical unemployment*. Milton Friedman, in his famous presidential address to the American Economic Association in 1967, coined the phrase “natural rate of unemployment” to refer to the rate that results from the equilibrium operation of the microeconomy, when macroeconomic conditions cause neither a general excess demand nor an excess supply of labor.

⁶Data taken from Chart 1.14 of OECD (1994), page 42 of Volume I.

At any point in time, macroeconomic conditions can lead to a slack aggregate labor market in which unemployment is above the natural rate or a tight labor market with unemployment lower than the natural rate. The difference between the actual rate and the natural rate of unemployment is often called cyclical unemployment. Keynes emphasized the importance of cyclical unemployment during the Great Depression, which he interpreted as a huge aggregate excess supply of labor.

When unemployment changes, there is often disagreement among economists about whether the causes are microeconomic or macroeconomic, in other words, whether it is a change in natural or cyclical unemployment. In the early postwar period, the natural rate of unemployment was widely regarded as being stable at about 4 percent in the United States and around 2 percent in Europe. Fluctuations in unemployment during the 1950s and 1960s were believed to be changes in cyclical unemployment around a fixed natural rate.

Changes in the labor market and in the general level of unemployment in the 1970s and 1980s convinced most macroeconomists that the natural rate can fluctuate considerably due to changes in the microeconomic structure of the labor market. The natural rate in the United States was reckoned to be 5.5 to 6.5 percent in the 1980s, but may have fallen somewhat in the 1990s. The causes of the high unemployment since the 1970s in Europe are regarded almost universally as microeconomic, which means that they should be viewed as increases in the natural rate.

Macroeconomists and labor economists have recently begun reexamining the sharp distinction between natural and cyclical unemployment. Despite the inconvenience it imposes on our theories, the microeconomy and macroeconomy are highly interdependent. A period of recession or depression caused by strictly macroeconomic factors will affect the microeconomic structure of the labor market in several ways. The demand for durable goods is usually more sensitive to business cycles than other goods, so these industries will shrink more than others in recessions, which will affect the industry, regional, and occupational structure of the demand for labor. On the supply side, workers who have been unemployed for a long time often lose job skills or job-finding skills, making them less likely to find a job.

Economists have used the term *hysteresis* to refer to situations where prolonged increases in cyclical unemployment raise the natural rate of unemployment. If hysteresis occurs, then the unemployment rate may never return all the way to its original natural rate after rising in a large and prolonged recession.

Within the category of natural unemployment, economists sometimes distinguish *frictional* and *structural* unemployment. Frictional unemployment results from the natural frictions of the labor-market matching process. You can think of the frictionally unemployed as job searchers for whom suitable vacancies exist, but who have not yet found these openings. Structural unemployment occurs when the skills and other characteristics of the unemployed do not match the requirements of the availa-

ble jobs. Technological change and structural shifts in the economy often cause changes in the skill composition of the job pool. If the labor force does not keep up with these changes, then structural unemployment is likely to result.

C. Introduction to Theories of Unemployment

There is no single unified model of unemployment. The Walrasian paradigm based on the market for a homogeneous good predicts that there should be no unemployment, so this workhorse benchmark model of neoclassical economics is not informative. Instead, one must move beyond the Walrasian model in one way or another. Since there are many ways in which actual labor markets differ from a Walrasian market, there are many possible approaches that can be followed.

While employment and unemployment are clearly connected in important ways, a theory of employment alone is not sufficient to explain the behavior of unemployment. One fallacy that is often committed by uninformed members of the public and the media is to assume that a decline in employment of, say, 1000 workers necessarily means that unemployment will rise by 1000. If a firm lays off 1000 workers, only a fraction will enter the ranks of the unemployed, and many of those are unlikely to remain there very long. Some of the laid-off workers will find jobs right away, moving from one employment position to another rather than into unemployment. Others will leave the labor force for retirement, education, parenting, or other non-labor activity. Similarly, when a firm hires 1000 new workers, some will have been previously unemployed but many others will come from other jobs or from outside the labor force as, for example, with new graduates finding their first jobs.

Rather than simply viewing unemployment as the counter-state to employment, we model it as a process of search. The success that individuals seeking new jobs will have in finding them depends on two broad kinds of circumstances: (1) the general balance of demand and supply in the labor market, and (2) the match between the searchers' characteristics and those of the available jobs. There are two broad categories of approaches to explaining movements in unemployment that correspond to these two kinds of circumstances.

One approach emphasizes the heterogeneity of workers and jobs. Because every worker and every job has unique characteristics, matching them up through a search process is time consuming. Search models examine the propensity of employers and job searchers to achieve matches and how that propensity varies over time. This approach models the flows of workers and jobs between states: a job match that results in a hire transforms an unemployed worker into an employed worker and a vacant job into an occupied one. To complete the model, one must examine the other labor-

market flows: job creation and destruction, entry to and exit from the labor force, and the flow of separations of existing workers from their jobs.

In the search approach, natural unemployment fluctuates when there are changes in the efficiency of matching in the economy or in the other flows between labor-market states. For example, if *structural shifts* in the economy make it more difficult to match the characteristics of unemployed workers with those of vacant jobs, then matching will be less efficient and the natural rate of unemployment will increase. A later section of this chapter looks at some of the literature on structural shifts and unemployment.

It should be emphasized that the kind of unemployment described by the search theories does not require a general excess supply of labor. It stresses the fact that even when the number of unemployed is equal to the number of job vacancies, neither number is likely to be zero.

The other major approach emphasizes microeconomic imperfections that lead to imbalance between demand and supply in the aggregate labor market, especially to excess labor supply. These imperfections can be associated with government interference such as minimum wages and unemployment benefits, or with deviations in the behavior of firms or workers from the assumptions of price-taking competitive markets such as the presence of unions or noncompetitive behavior by employers. This approach often tends to maintain the assumption that labor is a homogeneous good and emphasizes the possibility of a lasting imbalance between demand and supply in the unified labor market.

One popular model that takes that approach is the *efficiency-wage* model. This model is based on the notion that firms pay higher wages than would normally be necessary in order to attract workers. Different versions of efficiency-wage models stress different reasons why firms may do this. One is that higher wages may prevent shirking by employees; another is that a higher wage offer may attract a more qualified pool of applicants. While each firm in an efficiency-wage model wants to pay more than other firms, that obviously cannot happen. If all firms offer an efficiency wage, then aggregate wages in the economy are bid up above the market-clearing level and a general excess supply of labor results at the elevated wage.

Contract models are based on the observation that labor contracts often forbid firms from changing wages in the short run, but allow them to respond to variations in their need for labor through layoffs and overtime. A rich literature exists that examines the rationale for such contracts and their optimal structure.

Another model that follows this approach is the *insider-outsider model*, where a sharp distinction is drawn between the bargaining status of individuals who are currently working (insiders) and those who are unemployed or out of the labor force (outsiders). This model has been advanced as a promising explanation for the poor performance of labor markets in continental Europe in recent decades.

Romer discusses these three models in detail in Chapter 10. In addition to these, we will explore three other topics in unemployment in the next three sections of this chapter. In section D, we look at the connection between unemployment and minimum-wage laws. Section E examines how unemployment insurance affects job search and unemployment, and section F explores some possible ways in which the presence of unions might affect unemployment.

D. Minimum Wages and Unemployment

A lengthy empirical literature has examined the impact of minimum-wage laws on both employment and unemployment. Although there is a considerable range of results, the consensus seems to be that minimum wages have relatively minor employment/unemployment effects.⁷ Nonetheless, minimum-wage laws may have significant impacts in certain time periods and on certain groups of workers, especially teenagers and unskilled workers.

A simple minimum-wage model

The most basic analysis of the minimum wage proceeds in the way that we would analyze any price floor. In Figure 1, labor is assumed to be homogeneous; all workers participate in the same market and earn the same wage. If the market is Walrasian, then the wage will be w^* and the economy will operate with full employment at the level L^* . If a minimum wage is imposed at a level higher than the equilibrium wage, say at w_1 , then the market cannot reach equilibrium. Only L' workers are hired at w_1 and $L'' - L'$ workers are unemployed.

⁷A representative set of citations is Card and Krueger (1995), Deere, Murphy, and Welch (1995), Neumark and Wascher (1995), Brown, Gilroy, and Kohnen (1982), and Brown (1988).

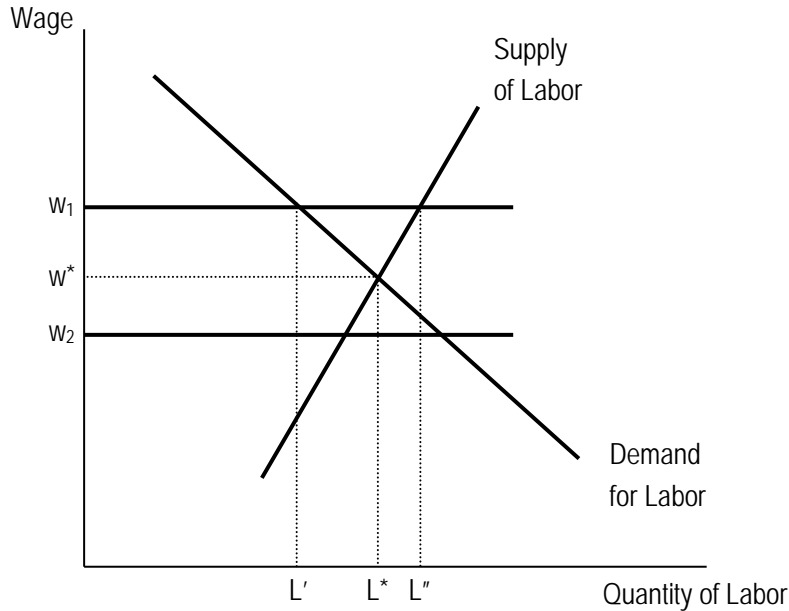


Figure 1. Minimum wage in unified labor market

Note that the effects of the minimum wage on employment and unemployment are not of the same magnitude. Employment falls from L^* to L' , while unemployment rises (by more) from zero to $L'' - L'$. This is because the higher wage draws additional workers into the labor market if the labor-supply curve slopes upward.

The magnitude of unemployment arising out of the minimum wage in this model depends on the elasticity of the labor-supply and labor-demand curves. If these curves are highly inelastic (steep), then the unemployment gap is small and the main impact of the minimum wage is to transfer income from consumers (through higher prices of products) or producers (through lower profits, if the product market is not competitive) to workers. If labor demand and supply are highly elastic (flat) then there will be a large reduction in employment and a large increase in unemployment.

Beyond its simplistic representation of the labor market, the analysis depicted in Figure 1 is flawed in one important respect: minimum wages are almost always set well below the average wage in the economy, not above it. Suppose that the minimum wage is set at w_2 in Figure 1. In this case, the minimum is not binding since no one is paying wages below w_2 in equilibrium. Thus, the minimum wage has no effect on the market at all. This is obviously too simplistic an analysis to capture how a minimum wage affects the market. Most workers, including virtually all skilled workers, earn more than the minimum wage and are not directly impacted by its presence. Some workers, mostly young and unskilled workers, may be affected,

however. To capture this kind of interaction, we need a segmented model of the labor market that separates skilled from unskilled labor.

Minimum-wage effects on skilled and unskilled labor

A two-sector labor market is shown in Figure 2. The left panel shows the equilibrium of the market for skilled labor. In this market, the equilibrium wage w_s exceeds the minimum wage w_m , so there is no direct effect of the minimum-wage law on unskilled labor. The right panel shows the unskilled labor market in which the equilibrium wage w_u is lower than the legal minimum. The wage floor is effective in the unskilled market, preventing demand from coming into equality with supply. As in our initial analysis of Figure 1, employment is reduced and an unemployment gap exists.

This would be the end of the story if there were no connections between the markets for skilled and unskilled labor. However, there may be spillovers on either the demand side or the supply side (or both). On the supply side, there would be no immediate spillover of workers from one market to the other. Unskilled workers cannot, presumably, become skilled immediately, while skilled workers earn a higher wage in the skilled market and have no incentive to move. In the longer run, supply flows in either direction are possible. Those who cannot find work in the unskilled sector due to the excess supply situation may choose to acquire skills and eventually move to the skilled sector. This would increase the supply of skilled workers and drive their wage down. However, the gap between skilled and unskilled wages has been reduced (for those unskilled who have work), so there may be less incentive for workers to acquire skills if they believe that they will be successful in getting an unskilled job at the higher minimum wage. This spillover would tend to offset the previous one, leaving the net effect on supply uncertain.

On the demand side, firms' demand for skilled workers may be affected by the increase in the wage for unskilled labor. If skilled and unskilled workers are substitutes, the firm will increase its demand for skilled workers, which will tend to push skilled wages upward. If they are complements, this will reduce skilled-labor demand and lower skilled wages. Although the substitute-complement relationship between skilled and unskilled labor is likely to vary across industries, the most common assumption is that they tend to be substitutes. If that assumption is true, then an increase in the minimum wage will raise the wages of skilled workers.

This hypothesis is supported strongly by the intense political support for minimum-wage legislation by labor unions. Most members of labor unions already earn more than the minimum wage, so they have no direct interest in a higher minimum wage. However, if firms substitute union workers for now-more-expensive lower-skilled workers, then a higher minimum wage for unskilled workers raises the whole

wage structure and union members may gain as well.⁸ Although union leaders may claim that their support of minimum-wage laws is philanthropy toward or solidarity with unskilled workers, it is highly unlikely that they would support these laws if they *reduced* the wages of union members.

To summarize, effective minimum-wage laws appear to benefit the fraction of unskilled workers that are able to find jobs. They reduce the welfare of those unskilled workers who cannot find employment. Skilled labor seems to gain from higher minimum wages as substitution by firms pushes the entire wage structure upward.

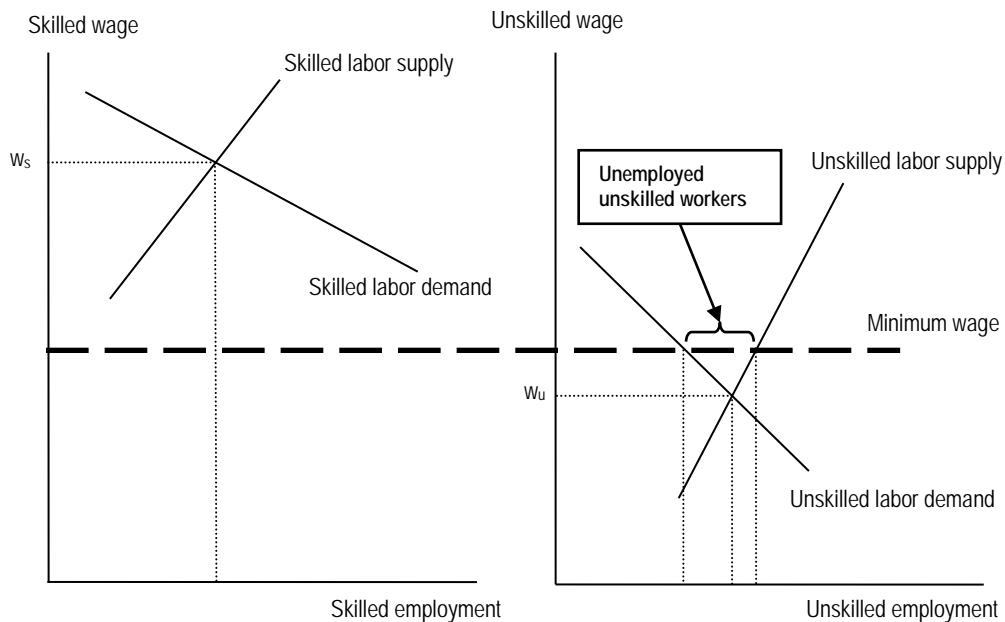


Figure 2. Minimum wage in a segmented labor market

A final word about the effects of minimum wages. In a perfectly competitive market, everyone who wanted to work would be able to find a job. But when jobs are rationed, as in the minimum-wage model, then some unskilled workers will find jobs and others will be unemployed. Who? What factors determine which unskilled workers will be the lucky ones? Any time that jobs (or anything else) is rationed by non-price means, the possibility of discrimination enters. Under competition, everyone finds a job and firms take what they can get. When there is excess supply at the prevailing wage, employers can pick and choose. In particular, teenage workers and members of recognizable minority groups may end up getting fewer of the available jobs if employers on average prefer to hire older and nonminority labor. Teens and

⁸Of course, higher general labor costs will eventually pass through to higher product prices, which will most likely eat away much of the gain in wages that skilled workers appear to get.

racial minority groups often have high unemployment rates; discrimination that arises under job rationing may provide a partial explanation.

E. Unemployment Insurance and Length of Job Search

A simple model of job search

One determinant of the rate of unemployment is the length of time that the average unemployed job-searcher takes to accept a new job. If searchers find and accept new jobs quickly, then unemployment is lower than if it takes a long time for people to move from unemployed to employed. The length of job search is sometimes modeled by considering the marginal costs and marginal benefits that a searcher expects from continuing to search. The search terminates when the marginal benefit of search no longer exceeds the marginal cost.

The principal cost of search is the wage income that is forgone by not having accepted the best offer received to date.⁹ The longer a worker searches, the better the job offers he or she accumulates, so the marginal cost of continuing becomes higher the longer is the search. The benefit of additional search is that a better job might be found. This marginal benefit is likely to decline as search continues, since the incremental increase in job quality is likely to become smaller as more jobs have been checked.

The declining marginal benefit curve in Figure 3 shows the falling marginal benefit of search, while the rising marginal cost curve represents the increasing marginal cost. Length of search is measured on the horizontal axis. Search equilibrium occurs where marginal benefit equals marginal cost, with duration D^* .

Anything that changes the marginal cost or marginal benefit of search will affect the chosen search duration, and thus affect equilibrium unemployment. For example, if a person's prospects for improving on his or her best job offer suddenly appear to improve, then the marginal benefit of search increases and the individual lengthens search time. If a searcher's spouse gets a better job, the marginal cost of search may fall, increasing the length of search. One possible explanation for why unemployment rose in the 1970s and 1980s when more and more married women entered the labor force is that two-income families may have lower marginal search costs

⁹ It may strike you as a big assumption to suppose that unemployed workers have job offers to refuse. The mass media encourage us to think of unemployed workers as having no choices, desperately in search of *any* job. While there are undoubtedly some unemployed who fit this profile, it is not typical. Most unemployed workers have the option of accepting a "poor quality" job such as working at McDonald's, even if they have received no offers in their usual occupation or at their accustomed salary.

than families with a single earner, allowing longer searches and raising the equilibrium rate of unemployment.

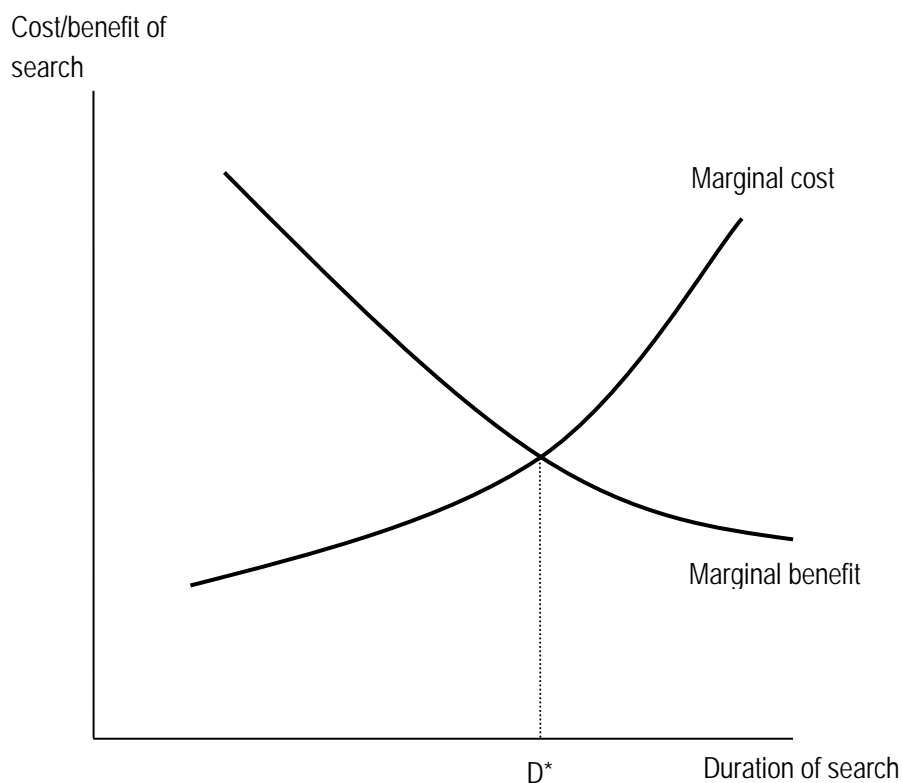


Figure 3. Marginal costs and benefits of search

Unemployment benefits and search duration

One consideration that has a large effect on search cost is the availability of unemployment-insurance benefits. If the government pays benefits to an unemployed worker to replace a share of his or her potential salary, marginal private search cost may be substantially reduced. In the United States, workers who lose their job are entitled to a share of their previous salary (usually about half, subject to an upper dollar limit) for a limited period of time (usually six months, though this is usually extended during recessions). The presence of limited-time unemployment benefits would shift the left-hand part of the marginal cost curve down as shown in Figure 4 assuming that benefits run out after a period of time equal to D' . The presence of unemployment benefits causes two potentially testable changes in unemployment. First, the duration of search and the equilibrium rate of unemployment should both increase. Second, because of the discontinuity in the marginal cost curve at the time limit for benefits, there should be an unusually large share of the unemployed accepting jobs at exactly that duration.

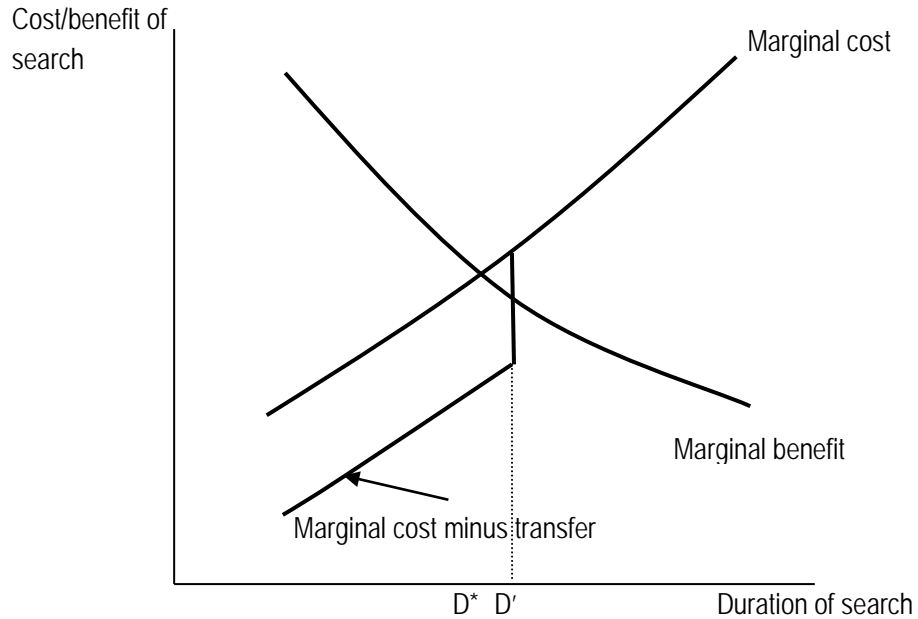


Figure 4. Equilibrium search with unemployment compensation

The effect of benefits on unemployment is difficult to test because of the many other factors that affect unemployment rates. In an empirical study of 1980s unemployment rates in 20 countries, Layard, Nickell, and Jackman (1991) find that an additional year of benefit eligibility cause unemployment to rise by 0.92 percentage points, and that an increase of one percentage point in the “replacement ratio” raises unemployment by 0.17 points.¹⁰ Time-series studies for the United States are less clear.¹¹ The effect of unemployment benefits on search can be seen more decisively from the effect of benefit duration on length of unemployment spells. A study by Bruce Meyer found a remarkable tendency for the length of unemployment spells to be exactly the maximum benefit duration. For example, if unemployed workers are eligible for six months of benefits, then an unusually large number of the unemployed would find jobs after exactly six months.¹²

¹⁰The replacement ratio is the percentage of previous income that the worker receives in benefits.

¹¹For a survey, see Atkinson and Micklewright (1991).

¹²See Meyer (1990).

Optimal search duration

Although there is considerable evidence that more generous unemployment benefits lead to higher equilibrium unemployment, it is not clear that this is necessarily bad. In order to assess the *optimal* duration of unemployment from a social point of view, we need to examine the marginal *social* costs and benefits of search, which may not be identical to marginal *private* costs and benefits.

Lengthening the search of an unemployed worker may lead to a better match between the skills of the worker and the requirements of the job. By using his or her skills in a better way, the worker receives higher wages. Society gains from this through improved productivity and efficiency—more output is available for society from the worker's effort. If wages reflect workers' marginal product accurately, then the social benefits of the increased search and improved matching correspond to the individual or private benefits.

The private cost to the worker of additional search is his or her forgone wages from the best available offer. Likewise, society loses the output that the worker would have produced had he or she not continued to search. If wages reflect marginal products then social costs and private costs will be similar.

If private benefits and costs of search match up with social benefits and costs, then the individual's choice of search duration will be socially optimal. If the government then introduces unemployment benefits to this situation, private search costs fall, but social costs do not change. Searchers will search longer than the socially optimal duration and the equilibrium or natural unemployment rate will be above the socially optimal rate.

However, there are reasons why we may be skeptical about the optimality of searchers' duration choices in the absence of unemployment benefits. One that is familiar from our analysis of consumption and investment is the possibility of liquidity constraints and imperfect capital markets. Households without substantial nonhuman assets usually find it difficult, and often impossible, to borrow at reasonable interest rates against future earnings. Consider the lone worker in such a household. If he or she were to become unemployed, the cost of a lengthy job search could be huge in terms of forgone utility (starving children come to mind). However, from society's point of view, the cost is merely the cash value of the worker's forgone wages. Whereas the worker cannot borrow against future earnings, society does so easily. Thus, the effective cost of search to the worker could be much larger than the social cost. In this case, the duration of search in the absence of unemployment benefits could be too short and the equilibrium unemployment rate too low. Introducing unemployment benefits may serve to offset (more or less) this externality, leading to a more socially efficient search decision.

F. Unions and Unemployment

The presence of a labor union that engages in collective bargaining on behalf of workers can alter the nature of the labor market in many ways. Several theories of unemployment incorporate the behavior of unions. In this section, we consider the direct effects of unions on employment and unemployment.

Whole libraries have been written about the goals, activities, and impacts of labor unions.¹³ While union activities and styles of organization differ substantially from country to country and over time, unions almost universally raise their members' wages through the process of collective bargaining. Unless unions also raise the marginal product of labor, this will result in a lower level of employment of union labor than would occur in a competitive market.¹⁴

An economy-wide labor union

The structure of labor unions varies considerably across economies. American unions are often defined by industry or by craft. For example, the United Auto Workers represents most workers in the automobile industry regardless of occupation, while individual craft unions represent workers in construction trades such as plumbers and electricians. Recent mergers and consolidations have often blurred industrial and occupations lines.

In some other OECD countries, unions are much more centralized. Finland, Norway, Australia, and Belgium have institutional regimes under which a single centralized wage bargain affects most of the entire economy. Austria, Denmark, and Sweden moved away from centralized wage bargaining in the 1980s. Most other countries in the OECD strike wage bargains that cover entire sectors of the economy. The United States, Canada, Japan, and the United Kingdom are exceptions in which wage bargaining at the enterprise or plant level predominates.¹⁵ This wide variation in the role of unions is reflected in a wide variety of models.

We first consider an economy in which wage bargaining is completely centralized with an economy-wide union. If the union bargains for a real wage higher than the equilibrium level, employment will fall along the labor-demand curve. Unem-

¹³ A good, but dated, survey of union economics is Freeman and Medoff (1984).

¹⁴ Some economists have argued that unions improve worker-firm communication and raise worker morale, which may cause productivity to increase. If unions increase productivity sufficiently, wages can rise without a decline in employment. This view is less popular than it was in earlier decades.

¹⁵ See the *OECD Job Study*, Volume 2, page 11, Table 5.9.

ployment will arise as workers seek to obtain union jobs but are unable to find employers willing to hire them at the union wage.

Although the model with an economy-wide union seems to be prone to high unemployment, that is not always the case. A union that represents the entire work force may recognize that a higher wage demand will reduce employment and that there is nowhere else for those workers to find work. This may induce such a union to moderate its wage demands somewhat. Through the 1970s and 1980s, the European economies that had centralized wage bargaining seemed to maintain lower unemployment rates than those with industry-level bargaining. However, this observation has been contradicted by rising unemployment in these countries in the 1990s.

A two-sector model of unions and unemployment

Since most countries have a substantial sector of the economy that is not covered by collective bargaining, it seems more realistic to model the labor market as having two sectors: a union sector and a nonunion sector.¹⁶ In the union sector, wages are above the equilibrium level due to effective bargaining. The nonunion sector functions competitively. For simplicity, the demand for labor in the two sectors is assumed to be similar, i.e., the marginal product of labor is the same in the union and nonunion sectors.

Suppose that we begin from an initial equilibrium in which wages are identical in the two sectors. When the union increases the wage through bargaining, firms will reduce employment of union labor. Thus, there may be some initial unemployment in the union sector and a differential between union and nonunion wages. Can this situation be sustained in equilibrium, or will demand and supply adjust further? As in the case of the minimum wage, there are possible spillovers between markets that may influence the ultimate equilibrium.

On the supply side, some union workers will be displaced by the reduction in union employment. What will these workers do? If they seek employment in the nonunion sector, then the supply of nonunion workers will increase and nonunion wages may fall. Thus, labor may spill over from the union to the nonunion sector. However, it is also possible that the attraction of high wages (if one is lucky enough to get a job) will retain and attract workers to the union sector. Union members and would-be-members may queue for union jobs, willingly accepting temporary or intermittent

¹⁶Although we will ignore the distinction in the text, unionization is not always synonymous with coverage by collective bargaining. The leading example of this is France, where only about 10% of workers belong to unions, but the collective bargaining agreements that unions negotiate extend to 90% of the work force. Rates of coverage by collective bargaining among OECD countries vary from over 90% in France and Belgium to about 20% in Japan and the United States. For simplicity, we refer to the part of the labor market characterized by collective bargaining as the union sector.

unemployment as a cost of getting higher wages during periods of employment. This queuing phenomenon—sometimes called *wait unemployment*—will raise the aggregate rate of unemployment.

In addition to these labor-supply spillovers, demand spillovers will occur if the goods produced by the union and nonunion sectors are fairly close substitutes. In this case, after a rise in the union wage, the nonunion sector can undercut the prices of firms in the union sector since it has lower labor costs. As the market price of goods falls, firms in the union sector may make losses and gradually leave the industry. The nonunion sector will absorb the extra demand, and the size of the nonunion sector will gradually grow relative to the union sector. Under this scenario, union power will be eroded over time, perhaps reducing the union wage differential and the share of the work force that is unionized.¹⁷ There will be little effect of unions on unemployment in the long run if demand spillovers are important.

Thus, a powerful but not universal union sector can cause wage differentials among industries and a decline in employment in covered industries. However, it is unlikely that very much of this decline in employment will result in unemployment. Only if workers queue for union jobs rather than spilling over into the nonunion sector will the existence of a union raise the equilibrium unemployment rate. Even this effect may diminish over time if substitution on the demand side reduces union power in the long run.

G. Sectoral Shifts and Unemployment

Central to the search theory of unemployment is the principle that workers and jobs are heterogeneous over many dimensions. Each worker has a particular package of skills, education, experience, location (and willingness to relocate), and preferences among job characteristics. Every job has particular characteristics of location, working conditions, and skill requirements. The process of search attempts to make matches between worker and job characteristics.

If the pool of searching workers and the pool of vacant jobs have broadly similar characteristics, then these matches will be made relatively easily and we would ex-

¹⁷It is tempting to suggest that this sort of substitution may account for the dramatic decline in U.S. unionization rates from about 30% in the 1950s to less than half of that today. However, the goods produced by union labor in the United States (typified by heavy industrial goods and craft trades) do not seem like close substitutes for those of the nonunion sector (such as agriculture and services). It seems instead that declining unionization has been a function of the general rise in service employment and a decline in manufacturing, largely motivated by changing demand, changing patterns of international trade, and shifts in technology.

pect short searches and low equilibrium unemployment. However, when there are fundamental differences between the characteristics of searchers and jobs, then matching will be more difficult and unemployment is likely to be higher. Thus, search theory suggests that the natural unemployment rate may be sensitive to changes in the efficiency of job matches. For example, if an economy is subject to an unusual degree of change in the sectoral composition of output, the occupational composition of the demand for labor, or the geographical distribution of output, then the natural rate of unemployment should be high.

The seminal study looking at the effects of sectoral shifts on unemployment is David Lilien (1982). Lilien argued that when the cross-industry dispersion of growth in employment is high, more workers will be changing industries and matching will be more difficult, leading to higher unemployment. He measured dispersion by the standard deviation across a broad set of industries of the rate of employment growth. When this variable was included in a regression in which the dependent variable was the unemployment rate, its sign was positive and statistically significant. Lilien's work was widely regarded as showing support for the idea (implicit in real business cycles) that a large part of cyclical fluctuations in unemployment can be explained by shifts in supply conditions rather than by aggregate demand.

The rationale for Lilien's result was challenged by Abraham and Katz (1986), who argued that his sectoral shift variable was capturing changes in aggregate demand rather than supply-induced structural shifts. Since the sectors of the economy have varying sensitivity to demand-induced business cycles, Abraham and Katz argued that such cycles would cause high cross-sectoral variability in employment growth. For example, investment and durable goods industries are more cyclically sensitive than services and nondurables, so when a (demand-induced) recession hits, employment will fall by more in these cyclically-sensitive industries than in others, which in turn causes the cross-sectoral variance in employment growth to increase.

Abraham and Katz supported their argument with evidence that job vacancies are negatively related to cross-sectoral variability.¹⁸ If the effect of high employment-growth dispersion was due to supply-induced sectoral shifts, then vacancies should be high in the growing industries and the unemployment rolls swollen with workers from the shrinking ones. Thus, Lilien's model predicts that unemployment and vacancies should both rise in times of high variability. Abraham and Katz argued that high measured variability occurs at times when overall labor demand is falling, which explains why vacancies are low rather than high.

¹⁸Since the United States does not collect data on job vacancies, the series they actually used was a measure of help-wanted advertising collect by the Conference Board, a private business organization. This variable is, at best, a rather weak proxy for job vacancies, which has prompted some proponents of the sectoral-shift model to question the validity of Abraham and Katz's results.

A lengthy literature has followed Lilien and Abraham and Katz, with significant evidence being presented on both sides of the debate.¹⁹ There is still considerable controversy about the relevance and interpretation of the association between unemployment and measures of sectoral shifts. Proponents of both sides (often Keynesians on the aggregate-demand side and neoclassical economists on the supply-induced sectoral-shift side) can find plenty of evidence to support their claims.

Another factor that should affect the ease with which searchers and job vacancies are matched up is the degree of flexibility that is present on both sides of the market. If workers are very narrowly trained for specific jobs in specific industries, then they should be much more prone to long unemployment spells if sectoral shifts arise. If, on the other hand, workers have a relatively general set of skills that are useful in a variety of industries and occupations (like Reed students!), then they should adapt quite easily to sectoral shifts.

H. Understanding Romer's Chapter 10

Romer examines three major categories of models: efficiency-wage models, contract models, and search and matching models. If time permits, we shall spend considerable time with each of these since all of them have strong influences on how modern macroeconomics think about unemployment. As noted above, the first two groups of models are commonly associated with a Keynesian view of unemployment and business cycles: imperfections in labor markets cause excess supply to be sustained. The search and matching models have found favor with those who prefer a more neoclassical approach.

Efficiency-wage models

Neoclassical factor demand theory treats the firm as a price taker in labor markets. The firm observes the equilibrium wage (for a particular category of labor) and determines how much it wishes to hire based on the marginal-revenue-product curve

¹⁹Significant papers in this debate include Loungani (1986; Loungani and Rogerson (1989), which argues that Lilien's sectoral shift variable moves almost entirely as a result of oil-price shocks; Loungani and Rogerson (1989), which uses micro data to examine worker transitions; Murphy and Topel (1987), which examines changes in unemployment and finds little support for the sectoral-shifts hypothesis, Davis (1987), which finds supportive evidence in the fact that sectoral shifts that reverse previous shifts seem to lower rather than raise unemployment; Blanchard and Diamond (1989), which looks in detail at the Beveridge curve, which relates unemployment to job vacancies; and (one of which I am especially fond) Parker (1992).

for that kind of labor. It has no control over its “wage structure” because it has no market power. It must pay the going wage in order to attract any workers; paying more than that amount is costly and pointless.

Or is it? The theory of *efficiency wages* is based on the premise that firms may voluntarily choose to pay more than the market-equilibrium wage for at least some categories of labor. Why would a firm choose to do this? In the strict neoclassical framework it would not, because the productivity of each worker is determined strictly by the technological factors that underlie the production function. Like robots, workers are plugged into the production process and are assumed to do their job regardless of wages, working conditions, or other factors.

In a richer model, worker productivity might be endogenously determined by such factors as their degree of effort or the rate of (disruptive) worker turnover. In such a setting, workers who recognize that they are receiving a premium above the “normal” wage may have incentive to work harder and may be less likely to quit. This raises productivity, perhaps by enough to more than compensate for the higher wage the firm pays.

To allow varying worker efficiency, we model the firm’s production as a function of eL , the effort of the average worker times the number of workers. Effort is assumed to depend positively on the real wage that the workers receive, $e = e(w)$, with $e' > 0$.

Note that there are many assumptions that could be used to justify this effort-wage or efficiency-wage relationship. For example, higher wages may make workers more content, healthier (if equilibrium wages in the economy are very low), more hard-working, less likely to cause problems for other workers, and less likely to quit. Any of these effects raises the worker’s productivity, so the efficiency-wage framework encompasses a fairly broad set of underlying assumptions.

To keep the model as simple as possible, we ignore capital. The firm’s real profits (measured in units of the output good) are $\pi = Y - wL = F[e(w)L] - wL$. The firm wishes to maximize this profit function. But what are its choice variables? We are used to firms that take w as given and choose the profit-maximizing level of L . However, in an efficiency-wage situation, firms are assumed to choose both the wage and the level of employment. Are there any constraints on these choices? Surely if the firm offers too low a wage it will be unable to attract as many workers as it wants.

Romer examines two cases. In the first case, there is unemployment in the economy and unemployed workers are assumed to be willing to work at any wage that the firm offers.²⁰ In the second, there are no unemployed workers and the firm must offer a wage at least as high as that of other firms.

To maximize profit in the unemployment case, the firm chooses the values of L and w that maximize the profit expression, Romer’s equation (10.4). The first-order

²⁰They have a zero “reservation wage,” in Economese.

conditions for this maximization are straightforward and lead to the condition (10.8). This condition can be interpreted as implying that the elasticity of effort with respect to the wage must equal one when the firm is at the optimal wage and employment level.

As Romer points out, the economy-wide equilibrium in this model can feature either zero unemployment (if aggregate labor demand at the efficiency wage is greater than or equal to supply) or positive unemployment (if labor demand is less than supply). Suppose that we are in a positive-unemployment equilibrium. What prevents firms from lowering the wage when unemployed workers come knocking at their doors offering to work for less than what the firm is currently paying? The answer is that they recognize that lowering wages will reduce employee productivity enough to raise labor costs by more than the saving in lower wages. Thus, the usual Walrasian mechanism that would tend to eliminate excess supply in the labor market is disabled in an efficiency-wage model and unemployment can remain indefinitely.

In section 10.3, Romer adds realism to the model by enriching the effort function to take account of factors other than the firm's wage that influence worker productivity. While the simple wage/efficiency link may be appropriate for a developing country in which higher wages enhance worker efficiency through better health and nutrition, the story is more complicated when the rationale behind the efficiency boost is effort, contentment, or retention. Even a firm that pays high wages may suffer from low morale if its wages are low relative to those of other firms in the market. A worker's contentment and level of effort probably depends more on her perception of her wage relative to other firms rather than on the actual level of the wage. Similarly, a higher rate of unemployment in the market is likely to cause workers to feel fortunate to have their current jobs and encourage them to work hard in order to retain them.

Thus, Romer augments the effort function to $e = e(w, w_a, u)$, where w_a is the wage paid by other firms in the market and u is the unemployment rate. Effort is increasing in the firm's own wage but decreasing in the wage of other firms. An increase in unemployment, other things being equal, raises effort. Since w_a and u are exogenous to the individual firm, the first-order conditions for setting the wage and employment level will not change, except that they will now depend explicitly on the alternative wage and unemployment rate.

In the Example sub-section of section 9.3, Romer assumes a functional form for the effort function given by (10.12) and (10.13). Equations (10.14) and (10.15) show the first-order conditions for this particular function. On page 465, he makes a key assumption: that all firms in the model are identical and that they will therefore all

end up choosing the same wage.²¹ Mathematically, this means that the aggregate equilibrium must have $w = w_a$, since if they were not equal then our chosen firm would be setting a different wage than that set by the other identical firms in the market.

Equation (10.17) shows that this model will generally result in positive unemployment. The intuition behind this result is straightforward. Each firm would like to pay an efficiency wage higher than its peers in order to elicit greater effort from its workers. In equilibrium, all firms pay the same wage, so no individual firm can encourage effort by paying more than others. However, as all firms bid up wages, the quantity of labor demanded falls below the quantity supplied, leaving some workers unemployed. The existence of this unemployment serves as a motivation to workers to work hard, since it increases the cost of losing their jobs. Thus, although each firm attempts to pay an efficiency wage in order to motivate workers to greater effort, it is the higher unemployment that results from these wage hikes that eventually serves that role.

Apart from the existence of equilibrium unemployment, the efficiency-wage model can explain one other observed anomaly. In a competitive labor market, given parameters values that are reasonable, we should see large fluctuations in real wages and only minor variations in employment when business cycles occur.²² In a market with efficiency wages, changes in the general level of product demand that cause fluctuations in labor demand may result in large changes in employment with relatively stable real wages. This seems more consistent with observed business cycles than the implications of the competitive, Walrasian model.

The Shapiro-Stiglitz model

Shapiro and Stiglitz (1984) entitled their paper “Equilibrium unemployment as a worker discipline device.” This idea clearly fits in well with the model of the previous section: high unemployment encourages greater effort. The Shapiro-Stiglitz model is a specific application of efficiency wages in which workers have an incentive to shirk (not work hard). Firms, obviously, would like to assure that workers instead exert high effort and thus achieve high productivity. If the firm can monitor worker performance at no cost, then it can simply insist on its desired level of effort as a condition of employment and fire workers who shirk. The level of performance required

²¹In game theory, an equilibrium in which everyone ends up making the same decision, given the decisions of all the other agents, is called a *symmetric Nash equilibrium*. We encountered this concept when we studied models of coordination failure.

²²Recall from our study of real business cycles that the high degree of observed fluctuation in employment was one aspect of the business cycle that the Walrasian model failed to explain.

and the wage would be set jointly at levels that would assure that the firm could attract workers.²³

However, the more interesting and realistic case is one in which firms cannot directly observe individual workers' effort. Instead, there is a given probability (less than one) that a shirking worker will be caught and fired. Thus, workers must decide whether to shirk or to work hard by balancing the increased utility of shirking against the probability of being caught and fired. This is, in itself, an interesting problem that warrants our attention. However, the worker's incentives depend in an important way on what happens to fired workers. If they can simply move immediately to another firm and shirk, then there is no real cost to being caught and no incentive to work hard. In order to motivate hard work, the fired shirker must lose something by being fired. In the spirit of the previous model, the fired worker either loses a wage premium (efficiency wage) offered by the employer or faces the prospect of an unemployment spell. Since all firms offer the same wage in equilibrium, it must be the existence of unemployment that gives workers an incentive to work hard. Hence the title "equilibrium unemployment as a worker discipline device."

Solving the Shapiro-Stiglitz model would be impossible without some simplification of our usual macroeconomic framework. For example, the instantaneous utility function (10.21) is obviously highly simplified. It neglects the tradeoff between consumption and leisure, has no diminishing marginal utility, and does not allow for variation in the marginal disutility of exerting effort relative to the utility of wage income (the coefficients on w and e are the same). However, for the purposes of the model, it is sufficient to focus attention on the key tradeoffs for workers. Similarly, the assumption that effort is discrete (taking on only the values 0 and \bar{e}) is obviously unrealistic, but introducing continuously varying degrees of effort makes the model much more complicated without producing any new insights.

As Romer notes, workers in the model are in one of three states. They may be employed and working hard (state E), in which case they have instantaneous utility equal to $w(t) - \bar{e}$. If they are in state S , employed but shirking (exerting zero effort), their utility is $w(t)$. Unemployed workers (state U) get utility of zero. The highest utility comes from being employed but shirking, but workers can remain in that state only until their shirking behavior is discovered, at which point they are fired and become unemployed.

The formal analysis of the Shapiro-Stiglitz model introduces us to the concept of *hazard rates*, which have become a popular tool for analysis of economic situations in which agents must predict when or if a particular event will occur. Hazard-rate mod-

²³ Since workers get utility from higher wages and lose utility from working hard, the combination of wages and work requirements offered by the firm must give workers a level of utility as high as that offered by other firms in order to attract workers.

els first evolved from the actuarial literature of the insurance industry (hence the name “hazard” that is attached). These models represent the probability that a particular event will occur during a period of time of given length by a hazard function such as the one shown in Romer’s equation (10.22). The hazard rate b is the instantaneous probability of changing from one state to another.

The Shapiro-Stiglitz model has three hazard rates; b is the probability that the job of an employed worker who is not shirking will end, q is the probability that a shirking worker will be caught and fired, and a is the probability that an unemployed worker will find a job.²⁴ These probabilities govern the transition of workers between states. The only decision that workers make in the model is whether to work or to shirk when they are employed. As in other efficiency-wage models, firms maximize profit by setting a wage level and employment level to maximize profit. The key difference between this model and the simpler efficiency-wage models is that the $e(w, w_u, u)$ function is endogenously determined: workers decide whether to work hard or shirk as a function of their wage and the unemployment rate. Similarly, the hiring rate a is determined endogenously by the balance of demand and supply in the aggregate labor market.

To solve the model, we first examine the decisions of firms and workers individually, then we close the model by looking at how these individual decisions interact in the labor market. Firms maximize profit at every moment as given by Romer’s equation (10.23); workers maximize lifetime utility from equation (10.20) and (10.21). The aggregate labor market enters because the job-finding rate a depends on the number unemployed.

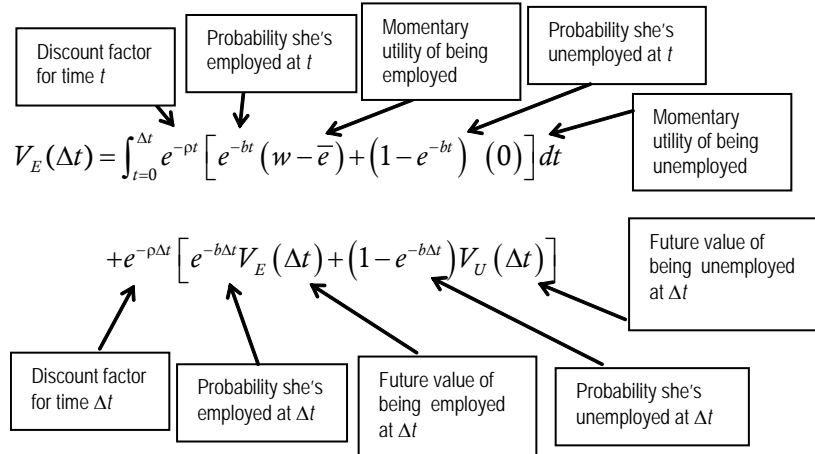
We begin by considering employed workers’ decisions about whether or not to shirk, i.e., about whether to be in state E or state S . If workers considered only the present, they would always choose to shirk, since that gives them the highest present utility. However, shirking raises the probability of being fired and experiencing a pe-

²⁴Units of measurement are a little tricky here. We usually think of probabilities as pure numbers, but that is not the case for hazard rates because the time dimension enters in an important way. For example, the probability that I will die is 1; the probability that I will die in the next year is less than one; the probability that I will die in the next week is less than that; and so on. Hazard rates, which measure the probability that an uncertain event will occur, must always be expressed per unit of time. In a discrete-time model, the obvious temporal metric is to say there is a probability of a that the event will occur and $(1 - a)$ that the event will not occur in the current period. In continuous time, we also measure the probabilities on a per-period basis, but since the event could happen at any instant within the period, the probability compounds like compound interest. That leads to the expression e^{-a} for the probability that the event does not occur at any moment during a period of length one. This formula is directly analogous to the formula used for continuously compounded interest in Chapter 4 of the coursebook.

riod of unemployment, which is the lowest-utility state. So workers must balance the present utility gain from shirking against the expected future utility loss from a greater probability of unemployment. Our basic decision rule is that workers will choose the state that gives them higher lifetime expected utility.

Romer represents by V_E the “value,” or expected utility, of working hard and by V_S the value of shirking. In order to evaluate the worker’s choices, we also need to know the value of being unemployed. The method by which Romer proceeds is known as *dynamic programming*. Our analysis consists of analyzing the behavior of an individual over a short interval of time Δt , then letting the length of the interval go to zero to get continuous-time behavior.

Consider Romer’s equation (10.24), which seems quite imposing upon initial inspection. The equation below shows how each term in this expression can be interpreted.



This expression measures the expected lifetime utility of a worker who works hard, broken down into the utility obtained over the interval from zero to Δt and the utility obtained after Δt . The first term (the integral) expresses the discounted utility that the worker gets between time zero and time Δt if she works hard. The second term in (10.24) measures utility after time Δt , given that she works hard.

One key distinction in equation (10.24) is between the V terms, which represent the lifetime expected utility of a person in who is currently in a particular state, and the terms $(w - \bar{e})$ and (0) , which represent the instantaneous utility one gets while in that state. The V terms are the capitalized lifetime (stock) value of the flow utility earned in the various state one may move through over one’s life. We integrate the instantaneous utilities in the first term to get total utility earned between 0 and Δt , then we add this to the lifetime expected utility of the state she ends up in.

The second term utilizes the V notation to express the value (expected utility) of being in the two states in which a hard worker might find herself. With probability $e^{-b\Delta t}$, the worker still has her job after Δt and the value at time 0 of being employed at

Δt is $e^{-\rho\Delta t} V_E$. Note that we must discount this value by $e^{-\rho\Delta t}$ in order to put it in terms comparable to time-zero utility. The worker finds herself unemployed at time Δt due to layoff with probability $1 - e^{-b\Delta t}$, and in that state she has lifetime prospects with period-zero utility of $e^{-\rho\Delta t} V_U$.

The integral in (10.24) is easy to evaluate. Because the exponential function is its own derivative (and integral),

$$\int A e^{Bt} dt = A \frac{1}{B} e^{Bt}.$$

In this case, A is $(w - \bar{e})$ and B is $-(\rho + b)$. That means that the definite integral in (10.24) evaluates to

$$\begin{aligned} \int_0^{\Delta t} e^{-(\rho+b)t} (w - \bar{e}) dt &= \left[-\frac{w - \bar{e}}{\rho + b} e^{-(\rho+b)t} \right]_{t=0}^{\Delta t} \\ &= \left[-\frac{w - \bar{e}}{\rho + b} e^{-(\rho+b)\Delta t} \right] - \left[-\frac{w - \bar{e}}{\rho + b} e^{-(\rho+b)0} \right] \\ &= \left(1 - e^{-(\rho+b)\Delta t} \right) \frac{w - \bar{e}}{\rho + b}. \end{aligned}$$

Plugging this expression into (10.24) gives equation (10.25), which has V_E on both sides. Solving yields equation (10.26). Taking the limit as Δt approaches zero gives the instantaneous value of being employed. This limit is somewhat tricky because the second term involves

$$\frac{1 - e^{-b\Delta t}}{1 - e^{-(\rho+b)\Delta t}}.$$

As $\Delta t \rightarrow 0$, all of the exponentials with Δt in the exponent approach one. That means that the expressions on the top and bottom of this fraction both approach zero. To find the limit, we use l'Hôpital's Rule, which says that the limit of such a fraction is the same as the limit of the derivative of the numerator divided by the derivative of the denominator. The derivative of the numerator with respect to Δt is $b e^{-b\Delta t}$ and the derivative of the denominator is $(\rho + b) e^{-(\rho + b)\Delta t}$. Taking the limit of this ratio as $\Delta t \rightarrow 0$ gives $b/(\rho + b)$, which leads to Romer's equation (10.27).

Note that (10.27) expresses the value of working hard as a function of the value of being unemployed. Following a very similar sequence of steps gives equation (10.29) as the lifetime expected utility of an employed person who shirks, and (10.30) as the lifetime expected utility of an unemployed worker.

Romer expresses these conditions in terms of the “return on a state,” which comprises the “dividend” earned while in the state, and the expected “capital gain or loss” from shifting from this state to another. The return is just the interest rate (or rate of time discount in this model) times the capital value of being in the state. The dividend is the utility yield from being in the state. The expected capital gain or loss is the amount that one expects to gain or lose from changing states in the future. Dynamic programming problems turn out to have first-order conditions that imply that the return on the state equals the sum of the dividend and the capital gain.

Since all workers are the same, either all will shirk or all will work hard. In order to get positive output, firms must pay a wage high enough to assure that no one shirks. Workers will work rather than shirk if $V_E \geq V_S$.²⁵ Since firms want to pay the least that they can while still assuring that no shirking occurs, they will pay a wage just high enough that $V_E = V_S$. Romer’s equation (10.35) is the level of the wage that satisfies this condition.

Substituting from the relationship between the job-finding rate a and the volume of employment L yields equation (10.37), which is the no-shirking condition in terms of the aggregate wage/employment locus. This is the upward-sloping curve in Romer’s Figure 10.2. As one moves to the right along this curve, the unemployment rate in the economy falls and the firms must pay a higher wage in order to prevent shirking. At a very high unemployment rate (to the left), only a low wage is required. If the firms in the economy offer a wage on or above the NSC curve, there will be no shirking in the economy.

Equilibrium in the labor market occurs where the no-shirking condition intersects the demand for labor curve, which results from the usual condition that the marginal product of labor equal the real wage. With this graphical apparatus, we are able to do a variety of comparative-static experiments to analyze the effects of changes in exogenous variables such as turnover and productivity as well as special cases of the model.

Implicit-contract models

The analysis of the implicit-contract model is much less complex than that of the Shapiro-Stiglitz model. The essence of the model is that workers are risk averse while the firm is assumed to be risk neutral. In this situation, it is mutually beneficial for the firm to offer implicit insurance to workers in the form of contracts that reduce the sensitivity of workers’ incomes in response to fluctuations in labor demand. In exchange for this insurance, workers accept lower wages than they would otherwise

²⁵Will employees work or shirk when $V_E = V_S$? In this borderline case, they are indifferent and may choose either. We will assume that they choose to work hard in the borderline case, since it makes the analysis easier. The results are affected only infinitesimally by changing this assumption.

demand. There is a mutual gain since the firm (which does not care about risk) raises its expected profit but incurs more risk, and the workers reduce risk enough to more than offset the utility lost from lower expected incomes.

The specific version of the model analyzed by Romer is simple because he assumes a finite set of possible states of the world A_i . With a discrete probability distribution over this finite set of alternatives, the expected value of profit in Romer's equation (10.40) and the expected value of utility in (10.42) are just summations across outcomes of the probability of each outcome times the profit or utility under that outcome.

The maximization of equation (10.43) is a standard Lagrangian in which firms maximize expected profits subject to workers' expected utility exceeding their reservation level of utility u_0 that they get from not working. The result (10.47) implies that firms will offer contracts that insure workers fully against income fluctuations.

Insider-outsider models

The insider-outsider model is a variation on the contract model that recognizes that the firm and a certain group of senior workers (continuing workers, perhaps) may conspire to improve their conditions jointly at the expense of less senior workers (new hires). Romer does not analyze an insider-outsider model in detail in the current edition of the text.

It is worth noting in passing that the labor-market structures of many European countries encourage this insider-outsider distinction. For example, to combat terribly high unemployment (over 20%) and low hiring rates, Spain passed laws allowing new workers to be hired under temporary arrangements, while existing workers could not be fired without incurring high separation costs.

The Blanchard and Summers model extends the idea of insiders and outsiders by allowing the number of insiders to change with shocks to productivity. In this situation, there will be no tendency for employment and unemployment to revert to "natural" levels. Employment in the economy will tend to follow a random walk, with the currently employed (insiders) protecting their employment status but reluctant to lower wages to bring in any outsiders.

Search and matching models

The search and matching model that Romer develops revolves around the *matching function* in equation (10.48). You can think of this as a Cobb-Douglas production function that "produces" matches from the "inputs" of unemployed workers and vacant jobs.

We are looking for a steady-state equilibrium at which the rate of unemployment is neither rising nor falling. That means that the expression for the change in employment equation (10.49) must be zero in equilibrium. There are two states that

workers can be in, employed and unemployed, and two states that jobs can be in, filled and vacant. The number of employed workers equals the number of filled jobs.

The solution is performed intuitively using the dynamic programming approach developed in the Shapiro-Stiglitz model. In each case, the “return” on being in a state is the interest rate times the value of being in the state. This is equated to the difference of the utility “dividend” earned while in the state minus the “capital loss” on the state, which is the expected value of the reduction in utility from changing states. This leads to equations (10.54) through (10.57).

The firm’s maximization condition is to set the level of vacancies in order to maximize profit. Since vacancies are assumed to cost nothing, this means that firms will expand vacancies until they have value zero. Solving the equations with the Cobb-Douglas matching function eventually leads to (10.70). This is graphed in Figure 10.6, and the comparative-static analyses follow using the graph.

I. Suggestions for Further Reading

On European unemployment

The OECD Jobs Study: Facts Analysis, Strategies, (Paris: OECD, 1994). (This is a summary of the detailed study.)

The OECD Jobs Study: Evidence and Explanations, (Paris: OECD, 1994), Volumes I and II. (These are the details.)

Layard, Richard, Stephen Nickell, and Richard Jackman, *The Unemployment Crisis*, (Oxford: Oxford University Press, 1994).

Unemployment: Choices for Europe, (London: Centre for Economic Policy Research, 1995).

Reducing Unemployment: Current Issues and Policy Options (Kansas City, Mo.: Federal Reserve Bank of Kansas City, 1994)

Theoretical models

Akerlof, George A., and Janet L. Yellen, eds., *Efficiency Wage Models of the Labor Market*, (Cambridge: Cambridge University Press, 1986).

Pissarides, Christopher A., *Equilibrium Unemployment Theory*, 2nd ed., (Cambridge, Mass.: MIT Press, 2000).

Blanchflower, Daniel G., and Andrew J. Oswald, *The Wage Curve*, Cambridge, Mass.: MIT Press, 1994).

Saint-Paul, Gilles, *Dual Labor Markets* (Cambridge, Mass.: MIT Press, 1996).

Lindbeck, Assar, and Dennis J. Snower, *The Insider-Outsider Theory of Employment and Unemployment*, (Cambridge, Mass.: MIT Press, 1988).

J. Works Cited in Text

- Abraham, Katharine G., and Lawrence F. Katz. 1986. Cyclical Unemployment: Sectoral Shifts or Aggregate Disturbances? *Journal of Political Economy* 94 (3):507-22.
- Atkinson, Anthony B., and John Micklewright. 1991. Unemployment Compensation and Labor Market Transitions: A Critical Review. *Journal of Economic Literature* 29 (4):1679-1727.
- Blanchard, Olivier Jean, and Peter Diamond. 1989. The Beveridge Curve. *Brookings Papers on Economic Activity* (1):1-60.
- Brown, Charles. 1988. Minimum Wage Laws: Are They Overrated? *Journal of Economic Perspectives* 2 (3):133-45.
- Brown, Charles, Curtis Gilroy, and Andrew Kohnen. 1982. The Effect of the Minimum Wage on Employment and Unemployment. *Journal of Economic Literature* 20 (3):487-528.
- Card, David Edward, and Alan B. Krueger. 1995. *Myth and Measurement: The New Economics of the Minimum Wage*. Princeton: Princeton University Press.
- Davis, Steven J. 1987. Fluctuations in the Pace of Labor Reallocation. *Carnegie Rochester Conference Series on Public Policy* 27:335-402.
- Deere, Donald, Kevin M. Murphy, and Finis Welch. 1995. Employment and the 1990-1991 Minimum-Wage Hike. *American Economic Review* 85 (2):232-237.
- Freeman, Richard, and James Medoff. 1984. *What Do Unions Do?* New York: Basic Books.
- Layard, Richard, Stephen Nickell, and Richard Jackman. 1991. *Unemployment: Macroeconomic Performance and the Labour Market*. Oxford: Oxford University Press.
- Lilien, David M. 1982. Sectoral Shifts and Cyclical Unemployment. *Journal of Political Economy* 90 (4):777-93.
- Loungani, Prakash. 1986. Oil Price Shocks and the Dispersion Hypothesis. *Review of Economics and Statistics* 68 (3):536-539.
- Loungani, Prakash, and Richard Rogerson. 1989. Cyclical Fluctuations and Sectoral Reallocation: Evidence from the PSID. *Journal of Monetary Economics* 23 (2):259-273.
- Meyer, Bruce D. 1990. Unemployment Insurance and Unemployment Spells. *Econometrica* 58 (4):757-782.
- Murphy, Kevin M., and Robert H. Topel. 1987. The Evolution of Unemployment in the United States: 1968-1985. *NBER Macroeconomics Annual* 2:11-68.
- Neumark, David, and William Wascher. 1995. Minimum-Wage Effects on School and Work Transitions of Teenagers. *American Economic Review* 85 (2):244-249.
- OECD. 1994. *The OECD Jobs Study*. Paris: OECD.
- Parker, Jeffrey. 1992. Structural Unemployment in the United States: The Effects of Interindustry and Interregional Dispersion. *Economic Inquiry* 30 (1):101-116.
- Shapiro, Carl, and Joseph E. Stiglitz. 1984. Equilibrium Unemployment as a Worker Discipline Device. *American Economic Review* 74 (3):433-44.