

UNEMPLOYMENT

WEEK 6

Fatih Kansoy *

University of Oxford

Hilary Term 2024-2025

*This is lecture notes for the course *Master in Financial Economics* at the Saïd Business School

Contents

1	Introduction and Terminology	4
1.0.1	Important Clarifications	5
1.1	Measuring Unemployment	5
1.1.1	Survey-Based Approaches	5
1.1.2	Measurement Challenges	6
1.2	Empirical Evidence on Unemployment	6
1.2.1	Time Series Evidence	6
1.2.2	Unemployment and Well-Being	7
1.2.3	Historical Evidence: The Great Depression	8
1.3	Labor Market Puzzles and Alternative Approaches	8
1.3.1	Determinants of Average Unemployment	9
1.3.2	Cyclical Behavior	9
1.4	The Game Plan: Alternative Theoretical Approaches	9
2	Efficiency Wages	9
2.1	Overview and Motivation	10
2.1.1	Potential Mechanisms Behind Efficiency Wages	10
2.2	The Basic Model	11
2.3	The Firm's Problem	11
2.3.1	First-Order Condition with Respect to L	12
2.3.2	First-Order Condition with Respect to w	12
2.4	Equilibrium Implications	12
2.5	Graphical Illustration of the Efficiency Wage Mechanism	13
2.6	Extensions and Critical Perspectives	13
2.7	Summary of Key Equations	14
2.8	Concluding Remarks	14
3	Extended Efficiency-Wage Framework and Case Studies	15
3.1	Recap of the Basic Efficiency-Wage Logic	15
3.2	Persistent Issues and the Need for a General Framework	16
3.3	A More General Efficiency-Wage Framework	16
3.3.1	An Illustrative Example: Summers (1988)	17
3.4	Case Studies: Integrating Historical and Modern Applications	18
3.4.1	Overview of Efficiency Wages in Practice	18
3.4.2	Empirical Case Studies: Ford and Costco	18
3.4.3	Comparative Analysis of Ford and Costco	19
3.4.4	Conclusions from the Case Studies	19
4	Extended Lecture Notes: The Shapiro–Stiglitz Model	20
4.1	Core Idea and Key Mechanism	20
4.2	Model Assumptions	21
4.2.1	Environment	21
4.2.2	Worker Effort and Utility	21
4.2.3	Poisson Processes and State Transitions	21
4.3	Firm's Problem and Profit Function	22

4.4	Dynamic Programming: Value Functions	22
4.4.1	Method 1: Infinitesimal Interval Approach	22
4.4.2	Method 2: Asset Equation Approach	23
4.5	No-Shirking Condition (NSC)	24
4.6	Aggregate Labor Market Equilibrium	24
4.6.1	Steady-State Flow for a	24
4.6.2	No-Shirking Condition in (L, w) -Space	25
4.6.3	Labor Demand	25
4.6.4	Graphical Depiction	25
4.7	Comparative Statics and Additional Insights	26
4.7.1	Impact of Better Monitoring ($q \uparrow$)	26
4.7.2	Impact of a Lower Turnover Rate ($b \rightarrow 0$)	26
4.7.3	Policy Implications and Efficiency Considerations	26
4.7.4	Cyclical Behavior: Layoffs versus Wage Cuts	27
4.8	Concluding Remarks	27
5	Introduction: Why Search & Matching?	28
5.1	Terminology and Basic Labor Market Concepts	28
5.2	Overview of the Search and Matching Framework	28
5.3	The Model Setup	29
5.3.1	The Basic Environment	29
5.3.2	Jobs and Firms	29
5.3.3	Matching Frictions and the Matching Function	30
5.4	Turnover, Unemployment, and the Beveridge Curve	30
5.4.1	Exogenous Separation	30
5.4.2	Steady-State Condition and the Beveridge Curve	31
5.5	Wage Determination through Bargaining	31
5.5.1	Surplus and Nash Bargaining	31
5.5.2	Dynamic Programming and Value Functions	32
5.5.3	Wage Determination via the Bargaining Condition	33
5.6	Vacancy Creation and the Job Creation Curve	33
5.7	Equilibrium and Comparative Statics	33
5.8	Graphical Illustration	34
5.9	Extensions and Research Directions	35
5.10	Concluding Remarks	36
5.11	Summary of Key Relationships	36

1. Introduction and Terminology

Unemployment persists even when the economy appears healthy, suggesting that real labor markets deviate from the ideal of frictionless, Walrasian equilibrium. In theory, every individual willing to work at the prevailing wage should be employed, yet persistent unemployment raises key questions: Why aren't all willing workers employed? Why do unemployment rates vary so widely across countries and over time (e.g., Japan under 5% versus over 10% elsewhere)? And what are the welfare implications—is unemployment a misallocation of resources or an unavoidable outcome of matching workers to suitable jobs?

A related puzzle is that while standard supply–demand theory predicts that a fall in labor demand should sharply reduce wages and keep employment nearly constant, real wages are sticky and employment fluctuates considerably. This indicates that factors such as wage-cutting barriers or matching frictions are important.

Two broad approaches have been developed. The traditional approach augments the basic supply–demand framework with mechanisms like *efficiency wages* (wages above the market-clearing level that boost productivity and reduce shirking) and *contracting models* (implicit contracts or long-term bargaining). The modern approach views the labor market as a collection of individual worker–job matches; in *search–and–matching* models, both workers and firms incur costs and delays in finding suitable matches, which naturally produces persistent unemployment even when vacancies exist.

In what follows, we review standard definitions and measurement issues. Unemployment is one of the most central topics in macroeconomics, and understanding its measurement is essential for both theoretical and policy analysis.

Employment is the number of individuals holding any form of paid work, whether full-time, part-time, salaried, or self-employed.

Unemployment refers to individuals who are not working but are actively seeking employment. According to the **ILO**, an individual is unemployed if:

- They are aged 15 or above (some countries use 16 or 18),
- They did not work during a specified reference period (typically one week),
- They are available to start work within the next two weeks,
- They have actively searched for work in the past four weeks (or have a job lined up to start soon).

1.0.1. Important Clarifications

- **Actively Seeking Work:** Activities include sending resumes, contacting employers, or using employment agencies.
- **Different Country Thresholds:** Although some countries use 16 or 18 as the minimum age, global standards typically set the threshold at 15 for comparability.

Labor Force: Denoted by Q , it is the sum of the employed and unemployed:

$$Q = (\text{Number of Employed}) + U.$$

Participation Rate: The ratio of the labor force to the total working-age population N :

$$\text{Participation Rate} = \frac{Q}{N}.$$

From these definitions, we derive:

$$\text{Unemployment Rate} = \frac{U}{Q}, \quad (1)$$

$$\text{Participation Rate} = \frac{Q}{N}, \quad (2)$$

$$\text{Employment-to-Population Ratio} = \frac{Q - U}{N}. \quad (3)$$

These measures together provide a comprehensive picture of labor market performance.

1.1. Measuring Unemployment

Accurate measurement is crucial but challenging. Most developed countries use large-scale household surveys to capture data on employment status, job search activity, and demographics.

1.1.1. Survey-Based Approaches

For example:

- In the United States, the **Current Population Survey (CPS)** interviews roughly 60,000 households monthly.
- In the United Kingdom, both the **Claimant Count** and the **Labour Force Survey** are used.

1.1.2. Measurement Challenges

Challenges include:

- **Discouraged Workers:** Those who stop searching are not counted, underestimating unemployment.
- **Survey Limitations:** Nonresponse, misreporting, and sampling errors can distort data.
- **Definition Discrepancies:** Variations in ILO guideline implementation complicate cross-country comparisons.
- **Underemployment:** Part-time or under-skilled workers may be employed but not fully utilized.

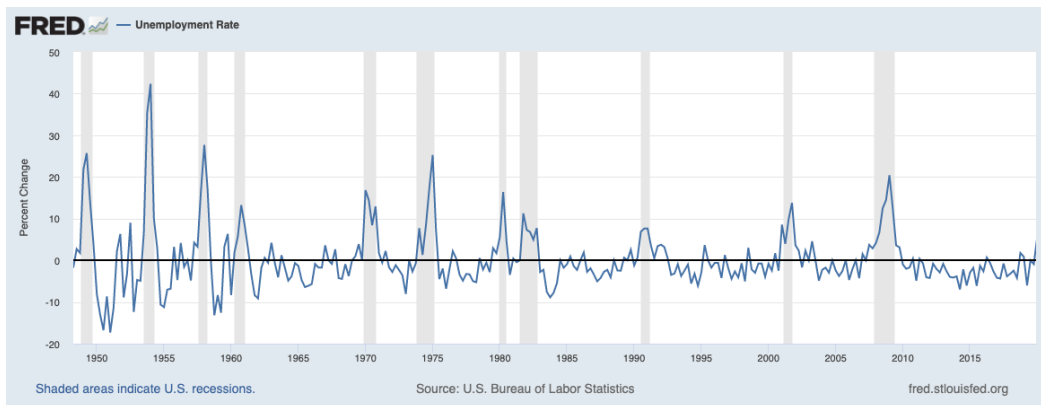
Analysts often supplement the official rate with broader measures (e.g., U-4 to U-6).

1.2. Empirical Evidence on Unemployment

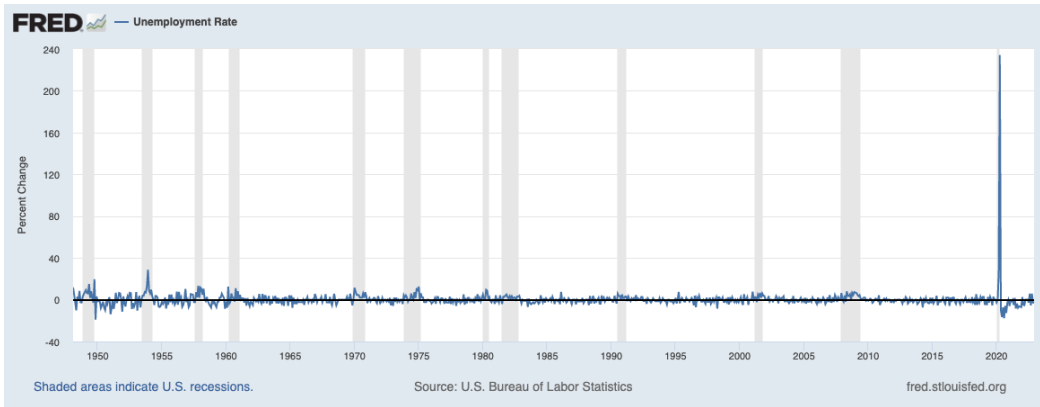
Empirical data reveal both cyclical and structural aspects:

1.2.1. Time Series Evidence

Time series data (e.g., Figures ?? and ??) show that U.S. unemployment rises during recessions and falls during expansions, but never reaches zero.



U.S. Unemployment Rate (1950–2019). Source: FRED, Series UNRATE.



U.S. Unemployment Rate (1950–2022). Source: FRED, Series UNRATE.

1.2.2. Unemployment and Well-Being

Studies (e.g., Winkelmann, 2014) find that unemployment is strongly negatively correlated with self-reported happiness. For example, a loss of employment is associated with a decline in subjective well-being comparable to major life events like divorce.

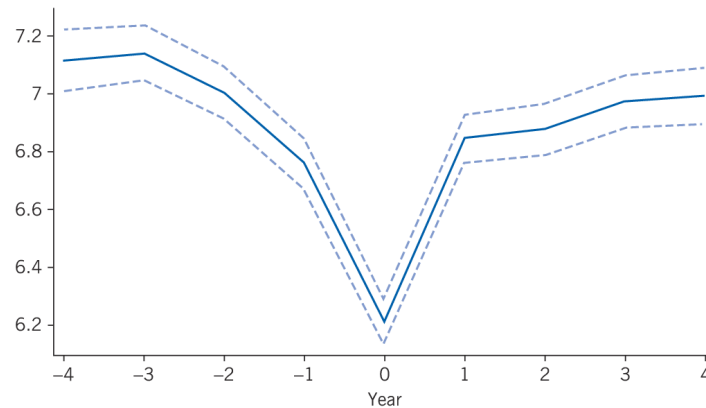


Figure 1. Relationship between unemployment and happiness. (Source: Winkelmann, IZA, 2014)

Data from Germany show a similar trend, where higher unemployment corresponds with lower reported happiness.



Figure 2. Evidence from Germany on unemployment and happiness. (Source: Winkelmann, IZA, 2014)

1.2.3. Historical Evidence: The Great Depression

During the 1930s, U.S. unemployment peaked at about 25%, with severe social and economic consequences. Iconic images from that era highlight the widespread hardship, influencing modern stabilization policies.



Figure 3. An image from the Great Depression illustrating the socioeconomic impact of 25% unemployment.

1.3. Labor Market Puzzles and Alternative Approaches

Empirical evidence reveals two major puzzles:

1.3.1. Determinants of Average Unemployment

Standard Walrasian models predict that wages adjust to clear the market, yet persistent unemployment suggests that institutional, cultural, or contractual factors (e.g., minimum wage laws, collective bargaining) prevent wages from falling sufficiently.

1.3.2. Cyclical Behavior

Wages tend to be sticky downward even during recessions, causing employment to adjust rather than wages. This behavior challenges models assuming full wage flexibility.

1.4. The Game Plan: Alternative Theoretical Approaches

Given these puzzles, economists have developed:

- a. **Traditional Approaches:** These extend the supply–demand model by incorporating mechanisms such as efficiency wages and implicit contracts to explain why wages do not fall enough to clear the market.
- b. **Modern Search-and-Matching Approaches:** These models view the labor market as a collection of decentralized matches between workers and firms, where search costs and frictions result in persistent unemployment even if wages are flexible.

Subsequent sections will explore efficiency wage theories, contracting models, and search–and–matching frameworks in detail.

In summary, this section establishes the key definitions and measurement challenges associated with unemployment, reviews empirical evidence on its cyclical and structural aspects, and outlines the theoretical puzzles that motivate alternative approaches to understanding labor market dynamics.

2. Efficiency Wages

Efficiency wage theories offer an alternative explanation for persistent unemployment by arguing that firms may *intentionally* pay wages above the market-clearing level. In doing so, firms boost worker productivity—whether by enhancing effort, reducing shirking, lowering turnover, or improving morale—even though this strategy may lead to involuntary unemployment when labor supply exceeds the firm’s desired employment.

2.1. Overview and Motivation

Definition – Efficiency Wage: An *efficiency wage* is a wage level set by a firm that exceeds the market-clearing rate in order to induce higher worker productivity (i.e., greater effort). Firms choose to pay this premium when the benefits—in terms of reduced shirking, lower turnover, or improved morale—outweigh the extra labor cost.

Motivation: Although standard Walrasian models predict that wages adjust downward to clear the labor market, empirical evidence reveals several persistent puzzles:

- **Relative Wage Rigidity:** Over the business cycle, employment fluctuates significantly while wages remain relatively stable.
- **Insufficient Wage Adjustments:** Even during periods of high unemployment, wages do not decline enough to eliminate the excess supply of labor.
- **Persistent Involuntary Unemployment:** Many individuals willing to work at the prevailing wage remain jobless, contradicting the prediction that unemployment is solely voluntary.

Efficiency wage models address these puzzles by suggesting that firms intentionally set wages above the competitive level in order to secure benefits in productivity, even if that choice results in a smaller labor force.

2.1.1. Potential Mechanisms Behind Efficiency Wages

Several mechanisms have been proposed in the literature to explain why paying a premium can yield productivity benefits that outweigh the higher wage cost:

- a. **Nutrition and Health:** In less developed economies, higher wages may enable workers to afford better nutrition and healthcare, thus enhancing their productivity. Although less central in advanced economies, this mechanism remains illustrative.
- b. **Reduced Shirking Under Imperfect Monitoring:** When it is costly or difficult for firms to monitor worker effort, a higher wage increases the opportunity cost of being fired, thereby discouraging shirking.
- c. **Attracting a Higher-Quality Applicant Pool:** When firms face asymmetric information about worker abilities, offering a higher wage can attract more skilled and capable workers, leading to improved overall productivity (cf. Weiss, 1980).
- d. **Fostering Loyalty and Reciprocity:** Workers who perceive their wage as fair or generous are more likely to exhibit higher effort and lower turnover, reinforcing the benefits of a

wage premium (as in Akerlof and Yellen, 1990).

Each of these mechanisms illustrates a situation in which the benefits of paying a premium wage exceed its costs, thereby justifying a wage above the market-clearing level—even if that means that some willing workers remain unemployed.

2.2. The Basic Model

Consider a representative firm that uses labor as its sole input. The production function is given by:

$$Y = F(eL),$$

where:

- L is the number of workers hired,
- e denotes the effort level per worker,
- eL is the effective labor input,
- $F(\cdot)$ is increasing ($F' > 0$) and concave ($F'' < 0$), ensuring diminishing marginal returns to effective labor.

Key Assumption – Effort Function: Worker effort is assumed to be an increasing function of the wage:

$$e = e(w), \quad \text{with } e'(w) > 0.$$

This assumption is central to efficiency wage theories: higher wages raise worker effort or reduce the incentive to shirk.

2.3. The Firm's Problem

The firm chooses both the wage w and the number of workers L to maximize its profit:

$$\pi = F(e(w)L) - wL.$$

Two scenarios are conceivable:

1. **Unconstrained Wage Setting:** In a labor market with a surplus of workers, the firm can set its wage freely, knowing that workers will choose to work for it if the wage is sufficiently high.
2. **Constrained Wage Setting:** In a tight labor market, the firm may be forced to pay at least the prevailing wage in order to attract workers, leading to an outcome closer to

competitive equilibrium.

In this model, we focus on the unconstrained case to illustrate how paying an above-market wage can arise.

2.3.1. First-Order Condition with Respect to L

Taking the derivative of the profit function with respect to L gives:

$$\frac{\partial \pi}{\partial L} = F'(e(w)L) \cdot e(w) - w = 0.$$

Thus,

$$F'(e(w)L) e(w) = w. \tag{4}$$

This condition equates the marginal product of effective labor to the wage.

2.3.2. First-Order Condition with Respect to w

Since effort $e(w)$ is a function of the wage, differentiating the profit function with respect to w yields:

$$\frac{\partial \pi}{\partial w} = F'(e(w)L) L e'(w) - L = 0.$$

Using the condition from (4), where $F'(e(w)L) = \frac{w}{e(w)}$, we have:

$$L e'(w) \frac{w}{e(w)} - L = 0 \implies \frac{w e'(w)}{e(w)} = 1. \tag{5}$$

This elasticity condition indicates that at the optimal wage w^* , a 1% increase in the wage leads to a 1% increase in worker effort, thereby minimizing the cost per unit of effective labor $\frac{w}{e(w)}$.

2.4. Equilibrium Implications

Let w^* and L^* denote the solutions to conditions (4) and (5). Two outcomes can emerge:

1. **Positive Unemployment:** If the aggregate labor supply exceeds the firm's optimal employment level L^* , some workers will remain unemployed. In this case, w^* is above the

competitive (market-clearing) wage, generating involuntary unemployment.

2. **Full Employment:** If labor demand is sufficiently high relative to supply, firms might have to offer wages above w^* to attract workers, potentially leading to full employment. However, efficiency wage models typically emphasize the persistence of unemployment.

Thus, by choosing w^* , firms maximize productivity by inducing the appropriate level of effort, even though this choice results in fewer hires overall—a trade-off that underlies wage rigidity and involuntary unemployment.

2.5. Graphical Illustration of the Efficiency Wage Mechanism

Figure 4 illustrates the determination of the efficiency wage. On the horizontal axis, the wage w is plotted, and on the vertical axis, the effective cost per unit of labor, $\frac{w}{e(w)}$, is displayed. Since $e(w)$ is increasing in w , there exists a wage w^* that minimizes $\frac{w}{e(w)}$. At w^* , the elasticity condition $\frac{w e'(w)}{e(w)} = 1$ holds, indicating that a proportional increase in w yields the same proportional increase in $e(w)$.

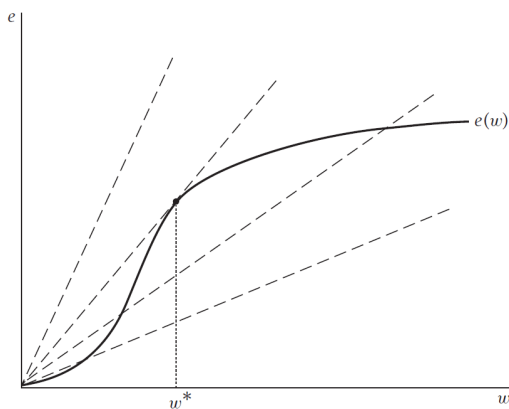


Figure 4. Determination of the efficiency wage by minimizing the effective labor cost $\frac{w}{e(w)}$. At the optimal wage w^* , a 1% increase in w leads to a 1% increase in $e(w)$.

2.6. Extensions and Critical Perspectives

While the basic efficiency wage model captures key features of wage rigidity and involuntary unemployment, it leaves open several important questions and areas for further exploration:

- **Alternative Compensation Schemes:** Might firms achieve similar benefits through non-wage forms of compensation, such as in-kind benefits (e.g., food or housing) or bonding

schemes where workers post a bond that is forfeited if they shirk? Such alternatives could potentially reduce administrative or legal costs.

- **Long-Run Dynamics:** The baseline model implies that if labor demand grows over the long run, unemployment should eventually vanish as wages rise. However, empirical evidence suggests that unemployment remains persistent. This calls for models that incorporate search frictions, evolving worker preferences, or institutional changes.
- **Role of Fairness and Social Norms:** Empirical studies (e.g., Akerlof and Yellen, 1990) indicate that workers' effort may also depend on perceptions of fairness, relative wages, and social norms. These behavioral factors can further reinforce wage rigidity.
- **Policy Implications:** If the wage premium is a private optimization by firms, then policies such as wage subsidies might alter firms' cost-benefit analyses. The net welfare effect of such policies depends on the trade-off between higher wages and the associated higher unemployment.

2.7. Summary of Key Equations

For reference, the central equations in the basic efficiency wage framework are:

1. Firm's Profit Function:

$$\pi = F(e(w)L) - wL.$$

2. First-Order Condition with Respect to L :

$$F'(e(w)L) e(w) = w. \quad (\text{Equation 4})$$

3. First-Order Condition with Respect to w :

$$\frac{w e'(w)}{e(w)} = 1. \quad (\text{Equation 5})$$

Together, these conditions determine the optimal (efficiency) wage w^* and the corresponding level of employment L^* . When the total labor supply exceeds L^* , the resulting gap produces involuntary unemployment.

2.8. Concluding Remarks

Efficiency wage models provide a powerful framework for understanding why wages may be rigid downward and why involuntary unemployment persists. By choosing to pay above the competitive wage, firms can boost worker effort, reduce turnover, and

enhance overall productivity, even though this comes at the cost of employing fewer workers. While the basic model captures these trade-offs through its production function, effort function, and first-order conditions, extensions that incorporate search frictions, behavioral factors, and institutional details are necessary to fully explain observed labor market dynamics. Nonetheless, the efficiency wage approach remains a central lens through which economists understand the persistence of wage rigidity and involuntary unemployment.

3. Extended Efficiency-Wage Framework and Case Studies

In this section we extend the basic efficiency-wage framework to address some empirical puzzles and illustrate its application with historical and contemporary examples. We first review the basic logic of efficiency wages, then discuss persistent issues that motivate a more general formulation, and finally combine the case studies of Henry Ford’s \$5-day policy and Costco’s high-wage strategy into a unified comparative analysis.

3.1. Recap of the Basic Efficiency-Wage Logic

Efficiency-wage models posit that firms may deliberately pay wages above the market-clearing level because higher wages induce workers to exert greater effort, thereby increasing productivity sufficiently to offset the higher wage cost. In the simplest version:

- A representative firm chooses wage w and employment L to maximize its profit:

$$\pi = F(e(w)L) - wL.$$

- Worker effort is modeled as an increasing function of the wage, $e = e(w)$ with $e'(w) > 0$.
- The firm’s optimization leads to the following first-order conditions:

$$F'(e(w)L) e(w) = w, \quad \frac{w e'(w)}{e(w)} = 1.$$

- The efficiency wage w^* is thus determined by the condition that a one-percent increase in w raises effort by one percent. Since w^* lies above the competitive wage, if the aggregate labor supply exceeds the firm’s optimal employment L^* , involuntary unemployment results.

3.2. Persistent Issues and the Need for a General Framework

While the basic efficiency-wage model explains why wages might be rigid and why firms choose to pay a premium, it also generates predictions that conflict with observed long-run trends:

- **Unchanged Real Wage Over Time:** The baseline model implies that the real wage remains constant until unemployment is eliminated, after which wages rise. In contrast, real wages tend to increase gradually with productivity.
- **Unemployment Trends:** The model predicts that, as labor demand grows, unemployment should eventually fall to zero. Empirical evidence, however, shows that unemployment often remains at a persistent, moderate level.
- **Short-Run versus Long-Run Dynamics:** In the short run, employment adjusts more than wages, consistent with wage rigidity; yet, over the long run, wages are expected to rise with productivity—a nuance not captured by the simple model.

These discrepancies motivate a more general efficiency-wage framework that enriches the basic effort function by incorporating external labor-market conditions.

3.3. A More General Efficiency-Wage Framework

To reconcile theory with empirical observations, we extend the effort function so that worker effort depends not only on the wage w offered by the current firm but also on external factors such as the wage w_a offered by other firms (the worker's outside option) and the overall unemployment rate u :

$$e = e(w, w_a, u).$$

This generalized effort function satisfies:

$$\frac{\partial e}{\partial w} > 0, \quad \frac{\partial e}{\partial w_a} < 0, \quad \frac{\partial e}{\partial u} > 0.$$

In other words:

- a. A higher wage w from the current firm increases effort.
- b. A higher wage w_a offered by other firms makes the current job less attractive, thereby reducing effort.
- c. A higher unemployment rate u increases the cost of job loss, encouraging greater effort.

Under these richer assumptions, the representative firm maximizes profit:

$$\pi = F(e(w, w_a, u)L) - wL,$$

taking w_a and u as given. The firm's first-order condition with respect to w generalizes to:

$$F'(e(w, w_a, u)L) = \frac{w}{e(w, w_a, u)},$$

and the optimal wage is determined by the generalized elasticity condition:

$$w \frac{\frac{\partial e}{\partial w}}{e(w, w_a, u)} = 1.$$

When all firms operate under these conditions, the economy-wide equilibrium determines w_a , u , and w , and if the wage required to sustain high effort remains above the competitive level, persistent involuntary unemployment ensues.

3.3.1. An Illustrative Example: Summers (1988)

A well-known variant by Summers (1988) specifies the effort function as:

$$e = \begin{cases} \left(\frac{w-x}{x}\right)^\beta & \text{if } w > x, \\ 0 & \text{otherwise,} \end{cases}$$

where

$$x = (1 - bu)w_a,$$

with $0 < \beta < 1$ and $b > 0$. Here, x serves as an index of labor-market conditions, incorporating the outside wage w_a and unemployment u . When $w \leq x$, workers exert no effort; when $w > x$, effort increases with the wage premium $(w - x)$ at an elasticity of β . Solving the generalized elasticity condition (i.e., setting the elasticity equal to 1) yields:

$$w = \frac{x}{1 - \beta} = \frac{(1 - bu)w_a}{1 - \beta}.$$

By imposing the equilibrium condition $w = w_a$, we obtain:

$$u = \frac{\beta}{b}.$$

Thus, in this extension, equilibrium unemployment depends solely on the parameters β and b , and it does not automatically vanish over time. This more general framework

is able to generate nontrivial, persistent unemployment levels along with a stable wage premium.

3.4. Case Studies: Integrating Historical and Modern Applications

To illustrate the theory in practice, we now consider two case studies that embody the principles of efficiency wages: Henry Ford’s \$5-day policy and Costco’s modern high-wage strategy.

3.4.1. Overview of Efficiency Wages in Practice

In the efficiency-wage model, a firm’s productivity is derived from its effective labor input $e(w)L$. By paying a wage w above the competitive level, a firm increases worker effort so that the effective cost per unit of labor,

$$\frac{w}{e(w)}$$

is minimized at the efficiency wage w^* . However, since w^* is above the market-clearing wage, if the labor supply exceeds the firm’s optimal employment level, involuntary unemployment results. This trade-off—higher wages inducing higher productivity but at the cost of employing fewer workers—is central to the efficiency-wage approach.

3.4.2. Empirical Case Studies: Ford and Costco

Henry Ford’s \$5-Day Policy: In 1914, Henry Ford famously doubled the daily wage of his assembly-line workers to \$5—a rate roughly twice the prevailing wage in the automobile industry. Ford described this policy as “one of the finest cost-cutting moves we ever made” (?). The higher wage reduced absenteeism and turnover while boosting worker morale and productivity. The increased effort, in turn, lowered the effective cost of labor (i.e., reduced the ratio $\frac{w}{e(w)}$). Although Ford’s premium wage led to some involuntary unemployment among workers who could not secure his high-paying jobs, the policy ultimately contributed to higher productivity and lower overall labor costs per unit output.

Costco’s Modern High-Wage Strategy: More recently, Costco Wholesale Corporation has adopted a similar strategy. In February 2025, Costco announced plans to raise hourly wages for most U.S. workers to over \$30 per hour over the next three years (??). For example, top-scale employees will receive an initial wage increase to \$30.20 per hour, with

planned incremental raises in subsequent years, while entry-level workers' wages will also rise. By consistently maintaining wages above the retail industry average, Costco aims to reduce turnover, enhance worker loyalty, and boost productivity—core predictions of the efficiency-wage model. However, the premium wage means that not all workers can be employed at these levels, contributing to some degree of involuntary unemployment.

3.4.3. Comparative Analysis of Ford and Costco

Shared Insights:

- Both Ford and Costco illustrate that paying a wage above the competitive level can reduce turnover and increase worker productivity. In both cases, the increase in effort (captured by the function $e(w)$) helps to offset the higher wage cost.
- The premium wage attracts a larger pool of applicants, leading to queues even though only a subset of workers can be employed at the higher wage.
- The resulting wage premium, while improving efficiency within firms, also contributes to involuntary unemployment when labor supply exceeds the firm's optimal employment.

Differences:

- Ford's policy was implemented in a manufacturing environment characterized by mass production and a moving assembly line, while Costco's strategy is deployed in the retail sector, where customer service and operational efficiency are paramount.
- The institutional and technological contexts differ significantly: Ford's early 20th-century labor market was less regulated, whereas Costco operates in a modern market with union influences, statutory minimum wages, and greater global competition.
- These differences imply that although the underlying efficiency-wage principles are similar, the magnitude, persistence, and broader consequences of the wage premium may vary across eras and sectors.

3.4.4. Conclusions from the Case Studies

Both historical and modern examples support the central claim of efficiency-wage theory: paying a wage premium can be optimal for firms because the resulting increase in worker effort reduces the effective labor cost. However, this strategy comes at the cost of employing fewer workers, leading to involuntary unemployment. Moreover, while efficiency wages explain key aspects of wage rigidity and labor market segmentation,

additional factors—such as search frictions, institutional differences, and technological change—must be integrated to fully account for observed labor market dynamics.

Overall, efficiency-wage theory remains a powerful lens through which to view wage determination and unemployment, highlighting the trade-offs between higher productivity and reduced employment. The experiences of Ford and Costco demonstrate that while the strategies may differ in context and implementation, the fundamental economic forces at work are similar.

4. Extended Lecture Notes: The Shapiro–Stiglitz Model

The Shapiro–Stiglitz model provides a microeconomic foundation for efficiency wage theories by explaining why firms may deliberately pay wages above the market-clearing level. In this framework, imperfect monitoring of worker effort means that firms must offer a wage premium to discourage shirking. This wage premium, in turn, prevents wages from falling to the level that would clear the labor market, thereby generating involuntary unemployment as a deliberate outcome—a mechanism often referred to as a worker discipline device or the no-shirking condition (NSC).

4.1. Core Idea and Key Mechanism

In many efficiency-wage models, it is assumed that higher wages lead to higher worker effort. However, Shapiro and Stiglitz (1984) provide an explanation for this relationship. Because firms cannot perfectly observe the effort levels of their employees in real time, workers face the temptation to shirk if the benefits of shirking outweigh the cost of losing their job. To counteract this incentive, firms offer wages above the market-clearing level, ensuring that the value of being employed is high enough to deter shirking. In equilibrium, all firms pay this premium wage, so that the labor supply exceeds the labor demand at the no-shirking wage.

- **Core Assumption:** Firms are unable to perfectly monitor worker effort; workers can either shirk (exert no effort) or work diligently (incurring an effort cost).
- **Key Mechanism:** If a worker shirks and is detected, she is fired. The threat of dismissal deters shirking only if being employed is sufficiently valuable—thus, wages must be set above the reservation or market-clearing level.
- **Result:** As a consequence, firms choose to pay above-market wages to avoid shirking, which leads to a situation where labor demand is lower than labor supply at that wage, generating involuntary unemployment.

4.2. Model Assumptions

4.2.1. Environment

- **Workers:** There is a continuum of identical workers (with measure \bar{L}) who discount future utility at rate $\rho > 0$.
- **Firms:** A large number N of identical firms operate in the economy, each maximizing instantaneous profit.
- **Timing:** The analysis is conducted in continuous time with a focus on the steady state.

4.2.2. Worker Effort and Utility

At any point in time, a worker can be in one of three states:

- (i) Employed and exerting effort (E), receiving instantaneous utility $w - \bar{e}$.
- (ii) Employed but shirking (S), receiving instantaneous utility w .
- (iii) Unemployed (U), receiving zero utility.

Here, $\bar{e} > 0$ represents the disutility or cost of exerting effort. The lifetime utility of a worker is given by:

$$U = \int_0^{\infty} e^{-\rho t} u(t) dt,$$

with the instantaneous utility $u(t)$ defined as:

$$u(t) = \begin{cases} w(t) - \bar{e}, & \text{if the worker exerts effort,} \\ w(t), & \text{if the worker shirks,} \\ 0, & \text{if the worker is unemployed.} \end{cases}$$

4.2.3. Poisson Processes and State Transitions

The model assumes that transitions between states follow Poisson processes:

- a. **Exogenous Job Separation (b):** Jobs end at an exogenous rate $b > 0$; the probability of a job surviving an additional time τ is $e^{-b\tau}$.
- b. **Shirking Detection (q):** A worker who shirks is caught at a rate $q > 0$. Thus, a shirker faces a combined separation rate of $b + q$.
- c. **Job-Finding (a):** An unemployed worker finds a job at rate $a > 0$. In steady state, a depends on both the number of vacancies and the number of unemployed workers.

4.3. Firm's Problem and Profit Function

Each firm produces output according to the production function:

$$\pi(t) = F(\bar{e}L(t)) - w(t) [L(t) + S(t)], \quad (1)$$

where:

- $F(\cdot)$ is increasing ($F' > 0$) and concave ($F'' < 0$), reflecting diminishing returns.
- $L(t)$ is the number of workers who are employed and exerting effort.
- $S(t)$ is the number of workers who are employed but shirking.
- The firm's objective is to maximize productivity by minimizing the number of shirkers; hence, it sets the wage w high enough to ensure that workers prefer exerting effort over shirking.

4.4. Dynamic Programming: Value Functions

A central feature of the Shapiro–Stiglitz model is the use of value functions to capture the expected discounted utility for workers in different states. The key value functions are:

$$V_E, \quad V_S, \quad V_U,$$

which represent the value of being employed and working, being employed but shirking, and being unemployed, respectively.

There are two standard approaches to deriving these value functions:

4.4.1. Method 1: Infinitesimal Interval Approach

In this approach, time is broken into a very small interval Δt . During this interval:

- The worker earns flow utility.
- There is a small probability of transitioning out of the current state due to either exogenous separation or, in the case of shirking, detection.
- Future payoffs are discounted.

Taking the limit as $\Delta t \rightarrow 0$ leads to the standard dynamic programming equations.

Value of Employment (V_E): For a worker exerting effort:

$$V_E(\Delta t) = \int_0^{\Delta t} e^{-(\rho+b)t} [w - \bar{e}] dt + e^{-\rho\Delta t} [e^{-b\Delta t} V_E(\Delta t) + (1 - e^{-b\Delta t}) V_U(\Delta t)].$$

Taking the limit as $\Delta t \rightarrow 0$ yields:

$$V_E = \frac{w - \bar{e}}{\rho + b} + \frac{b}{\rho + b} V_U.$$

Value of Shirking (V_S): For a worker who shirks:

$$V_S(\Delta t) = \int_0^{\Delta t} e^{-(\rho+b+q)t} w dt + e^{-\rho\Delta t} [e^{-(b+q)\Delta t} V_S(\Delta t) + (1 - e^{-(b+q)\Delta t}) V_U(\Delta t)],$$

which in the limit becomes:

$$V_S = \frac{w}{\rho + b + q} + \frac{b + q}{\rho + b + q} V_U.$$

Value of Unemployment (V_U): An unemployed worker receives zero flow utility but finds a job at rate a :

$$\rho V_U = a(V_E - V_U) \implies V_U = \frac{a}{\rho + a} V_E.$$

4.4.2. Method 2: Asset Equation Approach

Alternatively, each state is treated as an asset that yields an immediate dividend (its flow utility) and is subject to hazards that cause capital gains or losses. The value of the asset must satisfy:

$$\rho V_i = (\text{dividend}) - (\text{hazard rate}) \times (V_i - \text{value after transition}).$$

This approach quickly yields the same expressions for V_E , V_S , and V_U as obtained with the infinitesimal-interval method.

Bottom Line: Both approaches lead to the same dynamic programming equations for the value functions.

4.5. No-Shirking Condition (NSC)

A crucial aspect of the model is ensuring that workers choose to exert effort rather than shirk. To guarantee this, the wage must be set high enough so that:

$$V_E \geq V_S.$$

In equilibrium, firms set the wage so that the worker is just indifferent between exerting effort and shirking, i.e., $V_E = V_S$. Substituting the expressions from the dynamic programming equations yields:

$$V_E - V_U = \frac{\bar{e}}{q}.$$

This result implies that the *employment rent* (the difference between the value of being employed and being unemployed) must be at least \bar{e}/q . A lower detection rate q (i.e., harder monitoring) requires a higher wage premium to maintain the incentive to work. Combining this with the value functions, the firm's optimal wage is given by:

$$w = \bar{e} + (\rho + b + a) \frac{\bar{e}}{q}. \quad (5)$$

Thus, the wage exceeds the worker's disutility of effort by a margin that depends on the discount rate ρ , the separation rate b , the job-finding rate a , and the detection rate q .

4.6. Aggregate Labor Market Equilibrium

4.6.1. Steady-State Flow for a

Let L denote per-firm employment, so total employment is NL and total unemployment is $\bar{L} - NL$. With workers losing jobs at rate b and unemployed workers finding jobs at rate a , steady-state conditions require:

$$b(NL) = a(\bar{L} - NL),$$

which implies:

$$a = \frac{NLb}{\bar{L} - NL}. \quad (6)$$

A higher job-finding rate a reduces the cost of job loss for workers, thereby influencing the wage needed to deter shirking.

4.6.2. No-Shirking Condition in (L, w) -Space

Substituting expression (6) for a into (5) gives:

$$w = \bar{e} + \left(\rho + b + \frac{NLb}{\bar{L} - NL} \right) \frac{\bar{e}}{q}.$$

This relation shows that as the firm's employment L increases, the job-finding rate a rises, reducing the cost of job loss and requiring a higher wage w to maintain the no-shirking condition. Graphically, the NSC is upward sloping in the (L, w) -plane.

4.6.3. Labor Demand

The firm's labor demand is determined by the first-order condition:

$$\bar{e} F'(\bar{e}L^*) = w.$$

Because $F'(\cdot)$ is decreasing, the labor demand curve is downward sloping. Equilibrium is reached where labor demand intersects the NSC, and typically, $NL^* < \bar{L}$, which implies the existence of involuntary unemployment.

4.6.4. Graphical Depiction

Figure 5 schematically illustrates the equilibrium in the Shapiro–Stiglitz model. The diagram shows:

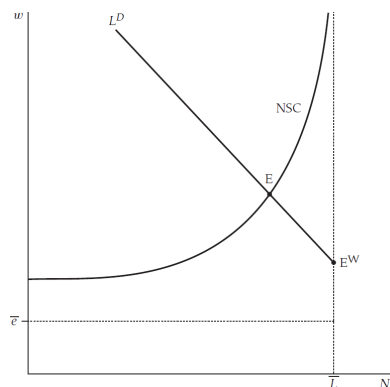


Figure 5. Equilibrium in the Shapiro–Stiglitz model. The NSC (No-Shirking Condition) curve and the labor demand (LD) curve intersect at point E , where the equilibrium wage exceeds the competitive level, leading to involuntary unemployment.

- The **NSC curve**, which is upward sloping in (L, w) -space.

- The **labor demand curve** (LD), which is downward sloping.
- Their intersection defines the equilibrium wage and employment level, with the difference between the labor force and L^* representing involuntary unemployment.

4.7. Comparative Statics and Additional Insights

4.7.1. Impact of Better Monitoring ($q \uparrow$)

When the detection rate q increases, monitoring improves, which reduces the wage premium required to deter shirking. Consequently, the NSC shifts downward, equilibrium wage falls, and employment increases. In the limit, as $q \rightarrow \infty$ (perfect monitoring), the model converges to the Walrasian outcome with full employment.

4.7.2. Impact of a Lower Turnover Rate ($b \rightarrow 0$)

If the exogenous separation rate b decreases to zero, once a worker is employed and exerts effort, she remains employed indefinitely. In this scenario, the only disincentive to shirk is the reduced probability of re-employment after being fired. The NSC becomes flatter, and the wage needed to deter shirking decreases. Unemployment may persist solely due to moral hazard considerations.

4.7.3. Policy Implications and Efficiency Considerations

The model suggests that decentralized equilibrium may be Pareto inefficient, as the wage premium necessary to deter shirking causes firms to hire fewer workers than would be socially optimal. Two potential policy remedies include:

- **Wage Subsidies:** Subsidies can lower the effective cost of paying a high wage, potentially allowing firms to employ more workers while maintaining the no-shirking condition.
- **Improved Monitoring:** Enhancing monitoring technology (increasing q) can reduce the required wage premium, thus mitigating the trade-off between productivity and employment.

These policy measures can help reduce inefficiency and lower involuntary unemployment.

4.7.4. Cyclical Behavior: Layoffs versus Wage Cuts

A central insight of the Shapiro–Stiglitz model is that firms cannot simply lower wages during downturns because doing so would weaken the no-shirking incentive. Instead, firms adjust by reducing employment—laying off workers—to restore the NSC. This mechanism explains why wages are relatively rigid downward even as unemployment fluctuates considerably over the business cycle.

4.8. Concluding Remarks

The Shapiro–Stiglitz model provides a microeconomic rationale for efficiency wages by demonstrating that, under imperfect monitoring, firms must pay a wage premium to deter shirking. This mechanism leads to an equilibrium in which wages are sticky downward and involuntary unemployment serves as a discipline device. Although the model assumes that workers are “rational cheaters” responding solely to monetary incentives, its core insight—that unemployment can be a strategic tool to maintain high effort—remains influential in modern labor and macroeconomic theory.

In summary, the model shows that:

- a. The no-shirking condition requires that the employment rent $V_E - V_U$ be at least \bar{e}/q , leading to a wage above the worker’s disutility of effort.
- b. The optimal wage is given by

$$w = \bar{e} + (\rho + b + a) \frac{\bar{e}}{q},$$

ensuring that the incentive to work is maintained.

- c. When combined with the firm’s labor demand, this wage premium leads to a situation where the labor market does not clear, resulting in involuntary unemployment.

This framework not only explains wage rigidity but also has important implications for policy—suggesting that measures to improve monitoring or reduce the required wage premium could help lower unemployment.

Finally, it is important to note that while the Shapiro–Stiglitz model emphasizes monetary incentives as the key determinant of worker effort, real-world behavior is also influenced by social norms, fairness considerations, and other non-monetary factors. Nonetheless, the model’s central contribution—that unemployment can serve a strategic role in deterring low effort—remains a cornerstone of modern labor economics.

5. Introduction: Why Search & Matching?

Traditional labor-market analysis often assumes that there exists a single market-clearing wage at which all individuals willing to work can immediately find a job. However, empirical evidence shows that unemployment persists—even when many vacancies are available. This raises the natural question: if jobs exist, why do some workers remain unemployed?

Search and matching models, most notably developed within the *Diamond–Mortensen–Pissarides* (DMP) framework, address this puzzle by explicitly modeling the frictions that arise when workers and firms search for each other. In these models, the labor market is viewed not as a single homogeneous pool but as a collection of heterogeneous workers and jobs that must be paired through a decentralized and costly search process. In what follows, we describe the core ideas, terminology, and mathematical structure of the search and matching framework.

5.1. Terminology and Basic Labor Market Concepts

Before examining the details of search and matching models, we review several standard definitions:

- **Working-Age Population, N :** The total number of individuals eligible to work.
- **Labor Force, Q :** The sum of employed and unemployed individuals.
- **Unemployment, U :** The number of individuals who are without a job but are actively seeking work.

From these, we obtain several key measures:

$$\text{Unemployment Rate} = \frac{U}{Q}, \quad \text{Participation Rate} = \frac{Q}{N}, \quad \text{Employment-to-Population Ratio} = \frac{Q - U}{N}.$$

While these aggregate indicators provide useful snapshots of labor market conditions, they do not capture the underlying heterogeneity or the dynamic process of search and matching.

5.2. Overview of the Search and Matching Framework

A central departure from classical, frictionless models is the view that the labor market consists of heterogeneous agents whose matching is not instantaneous. Instead:

- **Workers and Jobs are Heterogeneous:** Different workers have different skills and preferences, while jobs vary in their requirements and wage offers.
- **Decentralized, Pairwise Matching:** Rather than a centralized auctioneer who clears the market, workers and firms must actively search for one another. Workers submit applications and scan job postings, and firms post vacancies and screen candidates.
- **Costly and Time-Consuming Process:** The process of matching requires time and resources—advertising vacancies, waiting for responses, conducting interviews—which naturally leads to delays.
- **Bargaining over Wages:** Once a match is made, wages are not exogenously set by a market-clearing price but are determined by a decentralized bargaining process that splits the match surplus.

These features imply that even in a booming economy, some workers remain unemployed because the matching process itself takes time.

5.3. The Model Setup

5.3.1. The Basic Environment

- **Time:** Time is continuous, and our analysis focuses on the steady-state or long-run equilibrium.
- **Population:** The economy consists of a continuum of workers with total mass normalized to 1. Each worker is either employed (denoted E) or unemployed (U).
- **Preferences:** Agents are risk-neutral and discount future payoffs at a constant rate $r > 0$.

5.3.2. Jobs and Firms

Jobs (or firms) are modeled with the following characteristics:

- **Job States:** A job can either be **filled** (F) or **vacant** (V).
- **Output and Costs:** A filled job produces output $y > 0$ per unit time. If the worker is paid wage $w(t)$, then the firm's net flow payoff is $y - w(t) - c$, where $c > 0$ is a cost (for example, overhead or maintenance). A vacant job produces no output and incurs no wage costs but still costs c per period.
- **Profitability Condition:** We assume that $y > b + c$, where b is the unemployment benefit (or the worker's outside option), ensuring that a successful match yields a positive surplus.
- **Free Entry:** Firms may post vacancies as long as doing so is profitable. In equilibrium, the

value of a vacancy is driven to zero.

5.3.3. Matching Frictions and the Matching Function

The process by which unemployed workers and vacant jobs are paired is captured by the matching function:

$$M(U, V) = U m\left(\frac{V}{U}\right) = U m(\theta),$$

where labor market tightness is defined as:

$$\theta = \frac{V}{U}.$$

Empirical studies support the assumption that the matching function exhibits constant returns to scale, and that $m(\theta)$ is increasing in θ .

Job Finding and Vacancy Filling Rates:

- The **job-finding rate** for an unemployed worker is:

$$a = \frac{M(U, V)}{U} = m(\theta).$$

- The **vacancy-filling rate** for a firm is:

$$\alpha = \frac{M(U, V)}{V} = \frac{m(\theta)}{\theta}.$$

A higher θ means that vacancies are abundant relative to job seekers; thus, while the job-finding rate a is high, the vacancy filling rate α is low.

5.4. Turnover, Unemployment, and the Beveridge Curve

5.4.1. Exogenous Separation

Workers lose their jobs for reasons unrelated to the matching process. The exogenous separation rate λ captures this, so that the flow of workers exiting employment is $\lambda(1 - U)$, where $1 - U$ is the fraction of employed workers.

5.4.2. Steady-State Condition and the Beveridge Curve

In steady state, the flow into unemployment (due to separations) equals the flow out of unemployment (through job finding):

$$\lambda(1 - U) = aU = m(\theta)U.$$

Rearranging gives the steady-state unemployment rate:

$$U = \frac{\lambda}{\lambda + m(\theta)}.$$

This equation is known as the **Beveridge Curve**, which typically slopes downward in the (U, V) plane—indicating that higher matching efficiency (or higher θ) is associated with lower unemployment. The Beveridge Curve often shifts over time, reflecting changes in matching efficiency or structural features of the labor market.

5.5. Wage Determination through Bargaining

Once a worker and a firm are matched, they negotiate the wage according to a decentralized bargaining process.

5.5.1. Surplus and Nash Bargaining

The match between a worker and a firm generates a surplus. The worker's outside option is typically receiving the unemployment benefit b , while the firm's alternative is keeping a vacancy open (incurring cost c). The match surplus is then approximately:

$$\text{Surplus} = y - b - c.$$

Wage determination is modeled using Nash bargaining. Let $\phi \in (0, 1)$ denote the worker's bargaining power. The Nash bargaining problem is formulated as:

$$\max_w (V_E - V_U)^\phi (V_F - V_V)^{1-\phi},$$

where:

- V_E is the value of employment for the worker.
- V_U is the value of unemployment.

- V_F is the value to the firm of a filled job.
- V_V is the value of a vacancy (which, by free entry, is zero).

The outcome is that the worker obtains a fraction ϕ of the surplus and the firm obtains $1 - \phi$. Thus, the wage w tends to be an increasing function of the worker's outside option b , the job-finding rate a , the cost c , and the match output y .

5.5.2. Dynamic Programming and Value Functions

The determination of wages also depends on the dynamic programming problems solved by both workers and firms. The relevant value functions are defined as follows:

For Workers:

- *Employed Worker:* The value V_E satisfies

$$rV_E = w + \dot{V}_E - \lambda(V_E - V_U), \quad (6)$$

where \dot{V}_E is the time derivative (which is zero in steady state) and λ is the separation rate.

- *Unemployed Worker:* The value V_U satisfies

$$rV_U = b + a(V_E - V_U), \quad (7)$$

where b is the unemployment benefit and $a = m(\theta)$ is the job-finding rate.

For Firms:

- *Filled Job:* The value V_F satisfies

$$rV_F = (y - w - c) - \lambda(V_F - V_V), \quad (8)$$

where y is output, w is wage, c is the operating cost, and λ is the separation rate.

- *Vacant Job:* The value V_V satisfies

$$rV_V = -c + \alpha(V_F - V_V), \quad (9)$$

with $\alpha = \frac{m(\theta)}{\theta}$ being the vacancy filling rate. Under free entry, $V_V = 0$.

5.5.3. Wage Determination via the Bargaining Condition

The Nash bargaining solution for the match surplus is given by:

$$V_E - V_U = \phi [(V_E - V_U) + (V_F - V_V)]. \quad (10)$$

Since $V_V = 0$ under free entry, this condition, together with the value functions above, determines the wage w endogenously as a function of labor market tightness θ and other parameters such as $b, c, y, \lambda, r,$ and ϕ .

5.6. Vacancy Creation and the Job Creation Curve

Firms decide on the number of vacancies to post by comparing the cost of maintaining a vacancy with the expected benefit of filling it. The Bellman equation for a vacancy is:

$$rV_V = -c + \alpha (V_F - V_V).$$

By the free entry condition, $V_V = 0$, which implies:

$$V_F = \frac{c}{\alpha}.$$

Expressing V_F in terms of the net output (which depends on y and w) and the separation rate λ yields an implicit condition that determines the equilibrium value of labor market tightness θ . This condition is known as the **Job Creation Curve** (or Vacancy Supply Curve) and is often written as:

$$\theta \left(\phi + \frac{\lambda + r}{m(\theta)} \right) = \frac{1 - \phi}{c} (y - b - c). \quad (11)$$

5.7. Equilibrium and Comparative Statics

In equilibrium, the following conditions must simultaneously hold:

- a. **Labor Market Clearing (Beveridge Curve):**

$$U = \frac{\lambda}{\lambda + m(\theta)}.$$

- b. **Vacancy Supply (Job Creation) Condition:**

$$\theta \left(\phi + \frac{\lambda + r}{m(\theta)} \right) = \frac{1 - \phi}{c} (y - b - c).$$

c. **Wage Determination via Nash Bargaining:**

$$V_E - V_U = \phi [(V_E - V_U) + (V_F - V_V)].$$

These conditions jointly determine the steady-state values of unemployment U^* , vacancies V^* , labor market tightness $\theta^* = V^*/U^*$, and the wage w .

Comparative Statics:

- **Output Increase (y up):** A higher y increases the surplus from each match, encouraging firms to post more vacancies. This shifts the Job Creation Curve so that θ rises and, typically, unemployment falls.
- **Higher Separation Rate (λ up):** A higher λ increases the flow into unemployment, shifting the Beveridge Curve upward and raising steady-state unemployment.
- **Improved Matching Efficiency:** If $m(\theta)$ increases (due to better job search technology or training), then the job-finding rate increases and unemployment declines for any given number of vacancies.
- **Bargaining Power (ϕ):** An increase in worker bargaining power tends to raise wages, thereby affecting both the surplus split and the free entry condition, which in turn alters the equilibrium vacancy level.

5.8. Graphical Illustration

A schematic representation of equilibrium in search and matching models is provided in Figure 6. In the (U, V) -plane:

- The **Beveridge Curve** (depicted as a curved line) shows the inverse relationship between unemployment U and vacancies V in steady state.
- The **Job Creation Curve** (depicted as an upward-sloping line) represents the free entry condition for posting vacancies.
- Their intersection, denoted by point E , determines the equilibrium levels of unemployment and vacancies.

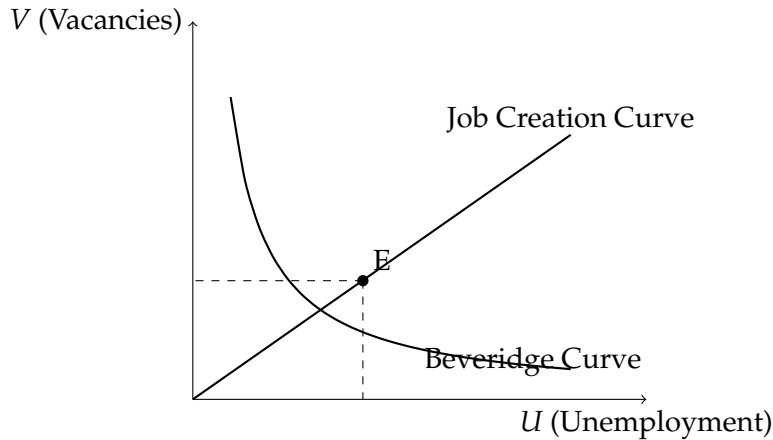


Figure 6. Schematic representation of equilibrium in search and matching models. The Beveridge Curve illustrates the inverse relationship between unemployment and vacancies, while the Job Creation Curve represents the free entry condition. Their intersection determines the steady-state equilibrium.

5.9. Extensions and Research Directions

Beyond the baseline framework, several extensions have been developed:

- **Worker Heterogeneity:** Incorporating differences in skills or job preferences can explain mismatches and partial matching.
- **On-the-Job Search:** Allowing employed workers to search for better opportunities introduces job-to-job flows and additional dynamics.
- **Directed Search and Wage Posting:** If firms publicly post wages, workers can direct their search toward more attractive opportunities, altering matching outcomes.
- **Incorporating Wage Rigidity:** Adding elements from efficiency wage theories or contractual arrangements can generate more realistic cyclical and wage rigidity.
- **Policy Implications:** Improvements in matching efficiency through better job boards, training programs, or relocation assistance can reduce frictional unemployment. Similarly, policies affecting separation rates or bargaining power can shift equilibrium outcomes.

Empirical evidence from large-scale data (e.g., Indeed, LinkedIn) supports the central role of labor market tightness $\theta = \frac{V}{U}$ in determining both job-finding and vacancy-filling rates.

5.10. Concluding Remarks

Search and matching models provide a compelling explanation for frictional unemployment by explicitly modeling the decentralized and time-consuming nature of matching workers with jobs. These models show that:

- Matching is a process that incurs search costs and delays, so even in an efficient market, some unemployment persists.
- Wages emerge through a decentralized bargaining process rather than being set by a central clearing mechanism.
- Firms post vacancies while weighing the costs of maintaining them against the benefits of filling them, resulting in a steady state where vacancies and unemployment coexist.

Although the baseline model does not fully capture all aspects of wage rigidity or cyclical fluctuations, it provides a critical foundation. In practice, additional factors—such as efficiency wage considerations, institutional constraints, and social norms—complement search frictions in shaping observed labor market outcomes.

5.11. Summary of Key Relationships

For clarity, we summarize the central relationships in the search and matching framework:

1. Matching Function:

$$M(U, V) = U m\left(\frac{V}{U}\right) = U m(\theta),$$

where labor market tightness is defined as $\theta = \frac{V}{U}$.

2. Job Finding Rate:

$$a = m(\theta).$$

3. Vacancy Filling Rate:

$$\alpha = \frac{m(\theta)}{\theta}.$$

4. Beveridge Curve:

$$U = \frac{\lambda}{\lambda + m(\theta)}.$$

5. Vacancy Supply (Job Creation) Condition:

$$\theta \left(\phi + \frac{\lambda + r}{m(\theta)} \right) = \frac{1 - \phi}{c} (y - b - c).$$

6. **Wage Determination via Nash Bargaining:** The wage is determined by the condition:

$$V_E - V_U = \phi [(V_E - V_U) + (V_F - V_V)],$$

where the value functions V_E , V_U , V_F , and V_V satisfy the corresponding Bellman equations:

$$rV_E = w + \dot{V}_E - \lambda(V_E - V_U), \quad (12)$$

$$rV_U = b + a(V_E - V_U), \quad (13)$$

$$rV_F = (y - w - c) - \lambda(V_F - V_V), \quad (14)$$

$$rV_V = -c + \alpha(V_F - V_V), \quad \text{with } V_V = 0. \quad (15)$$

Alternate Perspectives:

- *Micro Perspective:* Individual workers incur search costs and face probabilistic matching, which shapes their likelihood of finding a job and the quality of the match.
- *Macro Perspective:* The dynamic interplay between job creation (vacancy posting) and job destruction (separation) explains the persistence of unemployment and the relatively stable wage dynamics over the business cycle.
- *Policy Perspective:* Enhancing matching efficiency (through improved job search assistance or training programs) and adjusting policies affecting separation rates or bargaining power can have significant effects on equilibrium unemployment.