

## Fluctuations in Equilibrium Unemployment

By ROBERT E. HALL\*

Macroeconomists tend to play down the role of fluctuations in product demand in their accounts of the movements of employment and unemployment. In the reigning mode of thought, shifts in demand account only for transitory departures of the unemployment rate from its natural level. The more persistent movements of unemployment, which account for the bulk of its variance around a long-term mean, are assigned to changes in the natural rate itself. Current thinking about the natural rate puts its emphasis entirely on the supply side of the labor market. The recent decline of the U.S. unemployment rate to below 6 percent is seen as the consequence of the declining importance of young people in the labor force, just as the bulge in unemployment in the early 1970's was blamed on the baby-boom generation.

I offer the contrary view that shifts in product demand generate movements in employment and unemployment for more than a transitory period of disequilibrium. In this view, the determination of employment and unemployment takes place in an equilibrium setting where demand is on an equal footing with supply. However, the equilibrium is not necessarily the one achieved by perfectly competitive markets and strict profit maximization by the owners of firms. Both the level of unemployment and the amplitude of its fluctuations may be aggravated by certain features of the economy. In general, these features generate substantial wage differentials between vestibule workers and perma-

nent workers. For example, agency problems may prevent the owners of firms from operating their personnel policies so as to maximize the value accruing to owners. Instead, employees have an important role in hiring and firing, and are able to raise the value of long-term employment above the value of alternative activities. Or, limitations in worker-discipline devices may make it optimal for compensation to rise with experience faster than does productivity.

When permanent jobs pay substantially more than entry-level jobs, the rate of promotion is a critical determinant of the economic value derived by a worker taking a job at the entry level. In the model of this paper, movements of promotion rates resulting from variations in the rate of growth of permanent employment generate important variations in the shadow wage in the entry-level labor market, even though the actual cash wage is hardly variable. Changes in equilibrium unemployment are the result of positive elasticities of labor supply with respect to the shadow wage. To put it another way, unemployment rises in youth labor markets when opportunities for advancement to high-paying permanent jobs decline, even though there is no change in the immediate cash wage.

### I. Unemployment and Labor Supply

This paper focuses on the young adult labor market, involving workers aged from 18 to 30 when the transition from education and other nonwork activities into serious full-time work takes place. A good deal of unemployment arises from this group, and much of the growth of equilibrium unemployment in the past two decades has been concentrated in these ages. I do not attempt to deal with unemployment among younger

<sup>†</sup>*Discussants:* Olivier J. Blanchard, MIT; James Medoff, Harvard University; Robert Topel, University of Chicago.

\*Hoover Institution, Stanford University, Stanford, CA 94305.

teenagers seeking summer or after-school employment, or with the much lower unemployment among workers over age 30.

Much unemployment among young adults arises from their participation in activities other than career-related employment. The U.S. unemployment survey (the *Current Population Survey*) does not count an individual as unemployed if even one hour of work occurred in the survey week. On the other hand, people who are in school, spending time at home, or taking some time off are quite likely to be recorded as unemployed—any job-seeking activity, even just looking in the want ads, in the four weeks before the survey will classify a nonworker as unemployed. Hence, it seems reasonable to view unemployment in this group as primarily the complement of permanent employment. A substantial and stable fraction of young adults not started on career paths will be counted as unemployed. An explanation of variations in unemployment is virtually the same thing as an explanation of variations in employment. Times of high unemployment are times when a smaller fraction of young adults are started on career ladders. The fact that the great bulk of unemployment arises from individuals who work relatively little (Kim Clark and Lawrence Summers, 1979) supports this view of unemployment.

## II. A Model of Labor Supply Based on the Timing of Entry

The point of the simple model of this section is to show how the shadow or effective wage that drives labor supply takes account of the likelihood of promotion to higher-wage jobs. The model shows that promotion probabilities can be an important determinant of labor supply. The model focuses on the timing of entry to the labor force as the most important dimension of labor supply. To keep the model simple, I assume that activities before entry do not involve employment in the formal labor market, though in reality the activities may include part-time or episodic work. The model has the implication that entry occurs only once, which is another reasonable simplification, although it is not universally true.

Consider a young worker whose preferences about work are ordered by the utility function,

$$(1) \quad \int_t^{\infty} e^{-rs} y(s, t) ds + v(t).$$

Here  $t$  is the time of entry to the labor force,  $r$  is the discount rate,  $y(s, t)$  is expected earnings at time  $s$  given entry at time  $t$ , and  $v(t)$  is the utility derived from activities preceding entry.  $v(t)$  is an increasing, differentiable function satisfying  $v''/v' < -r$ . The first-order condition for the optimal time of entry is

$$(2) \quad -e^{-rt} y(t, t) + \int_t^{\infty} e^{-rs} \frac{\partial y(s, t)}{\partial t} ds + v'(t) = 0.$$

Define  $q(t)$  as the shadow value of employment or the effective wage:

$$(3) \quad q(t) = y(t, t) - \int_t^{\infty} e^{-r(s-t)} \frac{\partial y(s, t)}{\partial t} ds.$$

The effective wage comprises the immediate cash earnings of a new entrant,  $y(t, t)$ , plus the present discounted value of the subsequent advantage of earlier entry,  $-\partial y(s, t)/\partial t$ . If entering sooner makes promotion to a higher-wage job more likely to occur sooner, that is an extra value beyond the immediate cash wage.

In the steady state, where  $q(t)$  is constant over time, the first-order condition can be written compactly as

$$(4) \quad e^{rt} v'(t) = q.$$

Let the function  $\theta(q, r)$  be the optimal time of entry:

$$(5) \quad e^{r\theta(q, r)} v'(\theta(q, r)) \equiv q.$$

Under the assumptions stated, it is apparent that a higher effective wage brings earlier entry:  $\partial \theta(q, r)/\partial q < 0$ .

Now consider the labor force participation rate over the range  $0 \leq t \leq T$ , corresponding to ages 18 to 30. All those aged below  $\theta(r, q)$  will be out of the labor force and those above will be in the labor force. The labor force participation rate or labor supply function will be

$$(6) \quad n(q, r) = 1 - (\theta(q, r)/T).$$

Labor supply is an increasing function of the effective wage:  $\partial n(q, r)/\partial q > 0$ . A higher effective wage stimulates early entry and thus more total work from the age group. Unemployment is correspondingly lower. These changes can occur without any change in the current cash wage if the promotion rate increases.

As an example of the way that promotion rates influence the effective wage, consider the case where a worker starts at the entry wage  $w$  and faces a constant hazard,  $\pi$ , of promotion to a long-term job paying  $z$ . That is, after working for a length of time,  $\tau$ , the probability of being in the better job is  $1 - e^{-\pi\tau}$ . Then equation (3) shows that the effective wage is

$$(7) \quad q = \frac{r}{r + \pi}w + \frac{\pi}{r + \pi}z.$$

Figure 1 shows the effective wage as a function of the promotion rate for the case where  $r = 0.2$ ,  $w = \$5$ , and  $z = \$15$ . There is a substantial difference between the effective earnings of the entry job with little probability of promotion ( $\pi = 0$ ) and with a 30 percent annual probability of promotion ( $\pi = 0.36$ ).

It remains to establish a relation between promotion rates and the growth of a firm or an industry. Is it plausible that promotion rates fluctuate enough to cause important variations in the effective wage, given what is known about changes in growth rates of the sectors of the U.S. economy? Suppose that there are  $N_1$  vestibule workers and  $N_2$  permanent workers in a sector and let  $\alpha$  be the ratio of vestibule to permanent employment. Suppose that  $N_1$  and  $N_2$  and hence total employment are growing at the same rate,  $g$ . The flow of newly promoted workers is  $N_2$ . The number of vestibule workers is

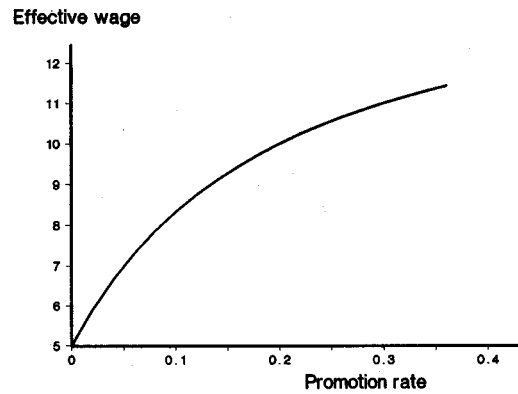


FIGURE 1. EFFECTIVE WAGE AS A FUNCTION OF THE PROMOTION RATE

$N_1$ . Hence the promotion rate is

$$(8) \quad \pi = \frac{\dot{N}_2}{N_1} = \frac{\dot{N}_2}{N_2} \frac{N_2}{N_1} = \frac{g}{\alpha}.$$

If vestibule workers are, say, one-tenth as numerous as permanent workers, then variations in the promotion rate are ten times as large as variations in the growth rate. Fluctuations of 1 or 2 percentage points in employment growth correspond to fluctuations of 10 or 20 percentage points in the promotion rate.

### III. Sources of Wage Differentials and Fluctuations in Promotion Rates

Economists have devoted great effort to studying and explaining wage differentials between beginning and senior workers. Not all explanations are consistent with a view that changes in promotion rates are a source of variations in the effective wage for starting workers. For example, there may be unobserved differences in productivity of workers at the time of hire. During the probationary period, firms accumulate information about productivity. They promote the workers who have a comparative advantage in working permanently for the firm. In that case, it is not easy to explain why promotion rates should vary with rates of growth. And even if promotion becomes more likely when

employment growth is high, the productivity of the promoted workers is likely to fall in relation to the entry-level workers, in which case the wage differential will probably shrink. An optimal promotion model seems unpromising as a way to explain important fluctuations in effective entry-level wages.

#### A. *Implicit Bonding*

One model, widely discussed in the literature on earnings profiles by seniority, explains wage differentials as a way to provide an appropriate penalty for workers who depart from jobs. The loss of future wages discourages workers from quitting and dissipating firm-specific human capital. The loss also makes threat of discharge an effective disciplinary device. The discharged senior worker must suffer some loss in earnings for the threat to be meaningful. In efficiency wage models, the loss occurs because the worker undergoes a period of unemployment before finding a new job. The alternative suggested here is that the more important wage loss is the period spent as a low-paid entry-level worker.

#### B. *Agency Problems in the Firm*

Another reason that firms might pay their permanent workers more than the market wage is that wage setting is imperfectly controlled by the owners of the firm. To some extent, employment decisions are made by employees, not by owners. The result is to divert some of the profit of the firm to employee-managers and away from the shareholders. The essence of the problem is the owners' inability to monitor their employees' efforts.

If the owners of a firm cannot observe their employees' efforts, it is desirable that they provide finance on a noncontingent basis. When the employee has no equity interest and the owner has a 100 percent equity interest, the employee is playing the game with someone else's money. Absent close monitoring, the performance of the employee is unlikely to be efficient. The answer is to make the employee play with his own money, by making the return paid to the owner be unrelated to the actual perfor-

mance of the firm. Thus employees are held to an exogenous standard of return on equity, or are required to pay a predetermined dividend. Whatever they make beyond that standard they are allowed to keep for themselves. The result will be earnings premiums in jobs where performance has exceeded the expectation formed when the system was established.

Now if one or a few of the workers are designated managers and can run the whole firm efficiently, the agency problem just set forth would have little impact on hiring and firing of the bulk of the work force. The managers would maximize profit by equating the marginal revenue product of labor to the market wage. But if a few managers could turn that trick, so could the owners. The agency problem exists throughout the firm. High-level managers are unable to monitor middle managers completely, who in turn are unable to monitor the heads of individual units. A firm is actually an association of worker-partners, none of whom is fully able to monitor the performance of the other worker-partners.

Figure 2 shows the standard analysis of the worker-managed firm. If all workers share equally in profit, and no worker outside the firm has a voice in employment decisions, employment will be set at the point  $L^*$  that maximizes the average revenue product of labor. At this point the marginal revenue product equals the average. The owner would choose the higher level of employment,  $\bar{L}$ , where the marginal revenue product of labor equals the market wage.

In practice, workers do not share equally in profits. Entry-level workers have nowhere near the claim of their senior brethren. By taking on workers in the vestibule, paid low wages, the partners holding the premium jobs can raise their own earnings above the maximum of the average revenue product of labor. They spread the profit created by all workers over the smaller number of worker-partners. However, if seniority is the only basis for excluding new hires from partnership, then the vestibule workers have to be let go before they become partners. Hence continual turnover of vestibule workers occurs. The turnover has the added benefit of permitting the partners to select only the

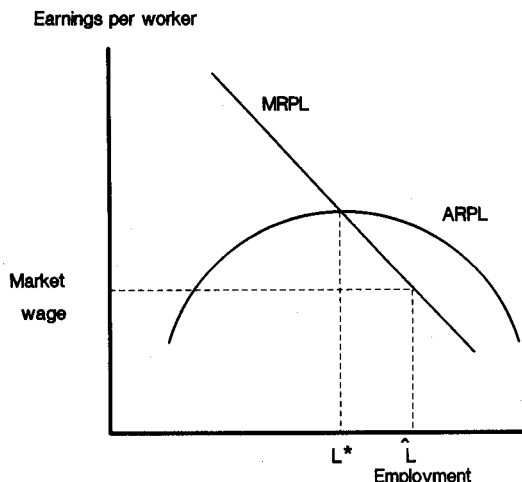


FIGURE 2. THE WORKER-MANAGED FIRM

most productive workers for partnership in case demand expands and it becomes attractive to add a long-term worker.

Because high-paid partners are recruited from among the vestibule workers, the cash wage is not the only attraction of vestibule employment. The total value is the cash wage plus the value associated with the probability of promotion. The value of the probability of promotion is the wedge separating the wage as perceived by the partnership and the wage as perceived by the worker. Promotion is much more likely when the firm is growing; the probability of promotion drops to close to zero when the number of partners is above the level that maximizes their average revenue product. The value of employment fluctuates much more than does the cash wage. Or, to put it the other way around, the firm finds it necessary to pay almost as high a cash wage in bad times as in good, because in bad times it cannot offer as high a chance of promotion. This situation is a consequence of the agency problem and the partnership rule that confers an instant capital gain to a new partner.

#### IV. Empirical Implications and Tests

In the labor market described above, there is no fixed natural rate of unemployment. Measured unemployment changes when incentives change. In times when vestibule jobs

offer high effective wages because promotion rates are high, young workers are attracted to those jobs away from other activities. Having a job is virtually a bar to classification as unemployed. Hence the move into jobs is accompanied by a decline in measured unemployment. In slack times, vestibule jobs are significantly less attractive even though they may pay the same cash wage. Unemployment is correspondingly higher.

The rate of growth of premium jobs is anything but a given constant over time. Consequently, the equilibrium rate of unemployment is variable over time. Periods of rapid growth of those industries with premium jobs will be periods of low equilibrium unemployment. When a shock occurs that is adverse to those industries, equilibrium unemployment can rise and remain high for many years. The theory permits highly persistent movements of unemployment; it does not require the return of unemployment to a constant normal level after a transient period of disequilibrium.

Measuring the growth rate of high-wage jobs is a relatively difficult task. Ideally, jobs paying premiums over wages elsewhere for workers with similar characteristics could be identified in cross-section research, and then the growth of employment measured in time-series data. What I have done to date is to observe that wage premiums are frequently found to be substantial in manufacturing in comparison to most other sectors. I have used the rate of growth of manufacturing employment as a rough measure of the opportunities available for promotion to advantageous long-term jobs. The postwar period has seen important changes in manufacturing employment growth over periods considerably longer than the business cycle. Growth was moderate in the 1950's, strong in the 1960's, slightly negative from 1970 until the early 1980's, and quite negative through the mid-1980's. In the work presented here, I take growth rates over 5-year periods to filter out some of the transitory business cycle effects.

With regard to unemployment, it appears possible to isolate the type of unemployment considered in this paper by taking the difference between total unemployment and insured unemployment. Insured unemploy-

ment is close to a trendless business cycle indicator; the notable increase in unemployment in the 1970's was almost entirely outside of insured unemployment (Gary Burtless, 1983). Uninsured unemployment is close to the concept of unemployment considered in the model of Section I, where unemployment arises among people who have not yet gained much job experience. The state-federal unemployment insurance system requires an extended period of employment in order to become eligible for very many weeks of unemployment compensation. As a result, few of the unemployed who have not yet started career employment receive benefits, even if their unemployment was preceded by a spell of employment.

Figure 3 shows the scatter diagram of manufacturing employment growth on the horizontal axis and uninsured unemployment on the vertical axis, measured as 5-year averages. There is a remarkably strong negative relation. The flow of job opportunities in manufacturing has declined almost continuously over the postwar period, except for the strong spurt of growth in the mid-1960's. The decline in manufacturing opportunities has coincided with an upward movement in the type of unemployment associated with delayed entry into serious career work.

Obviously the evidence in Figure 3 leaves many questions unanswered. Was there a similar decline in effective starting wages in sectors other than manufacturing? Is there more direct evidence other than uninsured unemployment about the changing behavior of young adults? But the evidence does suggest that changes on the demand side of the labor market have an important role in explaining changes over time in the natural rate of unemployment.

Summers (1986) considers a wide variety of evidence bearing on the issues examined in this paper. First, he shows quite convincingly that the upward trend in unemployment is unlikely to be the result of compositional shifts in the labor force. Second, he notes that unemployment has risen most among people who were previously employed, rather than new entrants to the labor force. Third, he notes that insured unemployment has not risen along with total un-

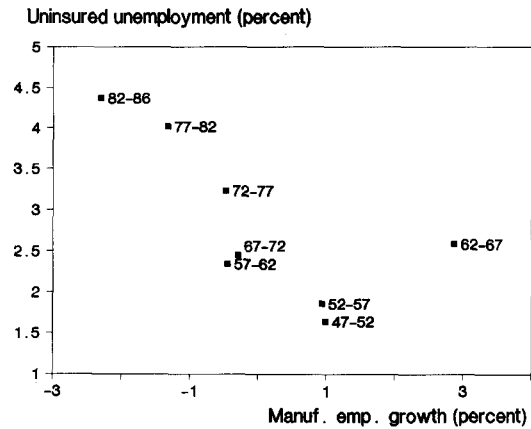


FIGURE 3. EMPLOYMENT GROWTH AND UNEMPLOYMENT

employment, a fact he considers a mystery. He is critical of the view taken here that changes in unemployment among those not yet fully committed to work are an important part of the explanation of the longer-term movements of unemployment. His primary evidence is that the rise in unemployment has not been concentrated among unmarried men. Ultimately, he concludes in favor of the view that more frequent and longer spells of unemployment between jobs are the principal source of higher unemployment. However, he does not develop a model in which this change is interpretable as an equilibrium. Moreover, Summers' inability to incorporate the rise in uninsured unemployment into his overall view is a significant shortcoming, in my view.

## V. Conclusions

When firms pay wage premiums to long-term workers, unemployment among young people is sensitive to the rate of growth of employment. The natural rate of unemployment is not fixed by supply considerations alone. Strong growth attracts entry-level workers because of the good prospects for promotion to long-term employment. In weak conditions, that factor is absent or attenuated and firms cannot take advantage of slackness by offering significantly lower cash wages. To put it differently, jobs that are attractive when demand is strong be-

cause they are on the first rung of a job ladder become dead-end jobs in slack times.

The effect of a decline in growth is anything but transient in this model. The period during which promotion to long-term jobs is closed off could easily be five years or a decade. The time pattern of regional differentials in unemployment fits this model reasonably well—they tend to come and go over near-decade periods. Similarly, as long-term demand shifts for durables occur, corresponding changes in employment and unemployment in that sector will occur.

## REFERENCES

- Burtless, Gary**, "Why Is Insured Unemployment So Low?," *Brookings Papers on Economic Activity*, 1:1983, 225-49.
- Clark, Kim B. and Summers, Lawrence H.**, "Labor Market Dynamics and Unemployment: A Reconsideration," *Brookings Papers on Economic Activity*, 1:1979, 13-60.
- Summers, Lawrence H.**, "Why is the Unemployment Rate So Very High near Full Employment?," *Brookings Papers on Economic Activity*, 2:1986, 339-83.