Part 2

Urban efficiency wages

In the first part of this book, we have developed different models of urban search matching. There is another dominant model in labor economics, which is the efficiency wage framework, initially developed by Shapiro and Stiglitz (1984).[20] The initial idea was to explain unemployment by high and rigid wages[21] in order to give some microfoundations for the Keynesian framework. Indeed, because shirking is very costly and monitoring is imperfect, firms set a sufficiently high wage to induce workers to work hard.[22] In equilibrium, wages are too high as compared to the market-clearing wages and downward rigid. As a result, unemployment emerges in equilibrium. In the search-matching framework, search frictions were responsible for unemployment while, here, high wages are the main culprit. In this part, we will expose the different urban efficiency wage models, starting with the benchmark models and then extending them in different directions.

Whether search frictions and bargain wages or efficiency wages are the correct model for explaining unemployment and wage setting is a matter of empirical relevance. We have seen that there is a host of empirical papers showing the importance of frictions in the labor market. What about efficiency wages?

The standard way of testing the efficiency wage model is very *indirect*. Indeed, the traditional ways of testing the efficiency wage theory is to examine whether there are large wage differences between sectors for identical workers that are unexplained by observable worker characteristics. This literature (Dickens and Katz, 1987; Kruger and Summers, 1988; Murphy and Topel, 1990; Neal, 1993) mainly shows that these unexplained differences can be attributed to differences in supervision/monitoring rates.

However, it is not straightforward to *directly* test the efficiency wage model, which is more generally known as the "rational cheater model". Indeed, the latter stipulates that employees are rational cheaters who anticipate the consequences of their actions and shirk when the marginal benefit exceeds the costs, and firms respond to this decision calculus by

---

[20]There are other versions of the efficiency wage models but here, we focus on the shirking version of Shapiro and Stiglitz (1984), since it is the most widely used.

[21]See the literature surveys by Akerlof and Yellen (1986) and Weiss (1991).

[22]In equilibrium, employed workers obtain "rents", which means that their (intertemporal) utility is strictly higher than that of the unemployed workers. Some researchers postulate that firms could design a contract that give not "rents" to workers (the so-called bonding critique, initially formulated by Carmichael, 1989). In that case, newly employed workers would post a "bond" that will be paid back to them later during their employment period. There are pros and cons on this debate and it is beyond the scope of this book to discuss the different possible contracts. The reader can consult the excellent book of Cahuc and Zylberberg (2004, Ch. 5) for an extensive discussion of this issue.

implementing monitoring and incentive pay policies (i.e. efficiency wage) that make shirking unprofitable. Cappelli and Chauvin (1991) have proposed a direct test of this model by looking at the relationship between the rates of employee discipline and the relative wage premium across plants of the same large firm. Their result suggests that higher wage premiums are associated with lower levels of shirking, as measured by disciplinary dismissals. Centering the empirical investigation on the effect of safety supervision by host employers on the wage of contract maintenance workers, Rebitzer (1995) also finds that high levels of supervision are indeed associated with lower wage levels. Using data from the US National Longitudinal Survey of Youth (NLSY) in 1992, Goldsmith, Veum, and Darity (2000) find that receiving an efficiency wage enhances an individual's effort and that individuals providing a greater effort earn higher wages.[23] However, as noted by Nagin, Rebitzer, Sanders, and Taylor (2002), there are two main problems in directly testing the shirking or rational cheater model. First, truly rational cheaters are more likely to engage in shirking behavior when such behavior is hard or expensive to detect. Second, should any association between monitoring and employee actions be found, it would be very difficult to disentangle the effects of monitoring strategies from responses to other unobserved features of the firm's human resource system. Resolving these problems requires a *natural experiment setting* in which monitoring levels are exogenously varied across similar sites and substantial resources are devoted to tracking the behavior of employees. Fehr, Kirchsteiger, and Riedl (1996) were the first to use experiments and show that higher wages indeed sharply reduce shirking. More recently, Nagin, Rebitzer, Sanders, and Taylor (2002) propose another experiment by collecting data from a large telephone solicitation company. The employees in this company work at 16 geographically dispersed sites. At each call center, telephone solicitors were paid according to the same incentive scheme, where the salary increased with the number of successful solicitations. This piece rate system, together with imperfect information about the outcomes of pledges, created incentives for employees to falsely claim that they had solicited a donation. To curb opportunistic behavior, the employer monitored for false donations by calling back a fraction of those who had responded positively to a solicitation. We are exactly in the framework of the shirking model of the efficiency wage. Nagin, Rebitzer, Sanders, and Taylor (2002) show that a significant fraction of employees behave according to the predictions

---

[23]See also Strobl and Walsh (2007) for a direct test of the efficiency wage between monitoring and effort.

of the shirking model. Specifically, they find that these employees respond to a reduction in the perceived cost of opportunistic behavior by increasing the rate at which they shirk.[24]

So far, we have seen that there is some empirical justification for the positive relationship between wages and effort, where the latter is not perfectly observable and thus, shirking is likely to arise. But another aspect of the Shapiro and Stiglitz (1984) model is that wage rigidity is the main explanation for (involuntary) unemployment. Indeed, in their model, firms refuse to reduce their wage even if some workers are ready to work for a lower wage because, in that case, the non-shirking constraint will not be met and thus, workers will shirk on the job. Different surveys of employers have shown that employers are indeed reluctant to reduce wages, even in the presence of high unemployment (Blinder and Choi, 1990; Agell and Lundborg, 1995; Campbell and Kamlani, 1997; Bewley, 1999). This is because employers believe that it would be unfair to employees to reduce their wages and that it would also have negative consequences in terms of work effort. Indeed, employers believe the main reason for avoiding pay cuts to be that they damage the morale and reduce productivity (Kaufman, 1982). In particular, firms are reluctant to hire workers who propose to work at a lower wage (Agell and Lundborg, 1995) because of the negative consequences it would have on workers' productivity. Fehr and Falk (1999) have studied the issue of downward wage rigidity in the framework of an experiment (in the laboratory) in which two groups, firms and workers, have the opportunity to agree on a wage contract through a mechanism of bilateral bidding. They show that under incomplete contracts, in which the effort level is not stipulated in advance, firms refuse to bid down the wage.

So, in the second part of this book, we are adopting this model to explain urban wages and unemployment.[25] This part, divided into three chapters, follows exactly the same structure as Part 1. In Chapter 4, we expose the simple models of the urban efficiency wage where the city is monocentric. We will first expose the benchmark urban efficiency wage model and then extend it to take into account more realistic situations such as endogenous housing consumption, open cities, resident landlords, long-run and free-entry of firms, different city-structures, etc. In each model of Chapter 1, we show how the urban wage is set by firms and

---

[24]There is a recent paper by Fehr and Goette (2007) which uses an experiment in a laboratory and shows that workers work more when wages are higher.

[25]To the best of our knowledge, the only test of the urban efficiency wage is due to Ross and Zenou (2008). Using the Public Use Microdata sample of the 2000 U.S. Decennial Census, their empirical results suggest that efficiency wages primarily operate for blue-collar workers, i.e. workers who tend to be in occupations that face higher levels of supervision. For this subset of workers, they find that longer commutes imply higher levels of unemployment and higher wages.

how urban unemployment is determined. We also calculate the equilibrium land rent in each location in the city. In Chapter 5, we provide different extensions of the benchmark model, for example, effort will vary with distance to the job center. As in Chapter 2, one important aspect that will be studied is when workers' relocation costs are no longer equal to zero but are very high. In that case, each change in employment status will not necessarily imply a change in residential location. Once more, this will have a profound impact on labor and land market outcomes. In the last Chapter of Part 2 (Chapter 6), we relax the assumption of monocentric cities and investigate different urban structures with multiple job centers. We show how each structure affects the labor market outcomes of workers as well as the determination of land rent in the city. One interesting issue will be the study of rural-urban and urban-urban migration.