Chapter 11
Pay and Productivity:
Wage Determination within the Firm

Summary

Beginning with the overview of the labor market presented in Chapter 2, the organizing theme of Modern Labor Economics has been the demand and supply model of the labor market. Chapters 3 to 5 focused primarily on modeling demand decisions, while Chapters 6 to 10 emphasized supply. Understanding demand and supply decisions is essential for anyone wishing to think seriously and critically about labor issues. The remaining chapters in the text take up a series of special topics that build on the foundation provided by the modeling of demand and supply decisions to form a more complete framework for thinking about labor issues.

The first special topic taken up in Modern Labor Economics is the complex relationship between compensation and productivity. Although the early models of the firm’s labor demand decision assumed that the firm simply took the market wage as a given when deciding on employment, later chapters suggested that the firm’s compensation decision may be more involved than that. In Chapter 5, for example, the firm had to decide on a wage that would allow it to recover its specific training costs and at the same time keep the worker from quitting. In Chapter 8, the firm had to decide on the mix of wages and employee benefits that would allow it to attract the workers needed to stay competitive. In this chapter, additional decisions relating to the basis, level, and timing of compensation are explored. These decisions, along with the ones discussed in earlier chapters, come together to form the firm’s compensation policy.

Managers must take into account that there are many factors that influence worker productivity, other than the human capital and physical capital issues discussed previously. Workers differ from each other in work habits, which cannot generally be observed prior to hiring. A given worker’s productivity, all other things equal, may vary over time, in different environments, or according to the worker’s motivation. A worker’s productivity is a function of innate ability, level of effort, and factors relating to the worker’s environment; being very productive is sometimes or usually related to taking initiative rather than simply following orders. Thus compensation should be designed to attract the right type of employees and give them the right incentives to be as productive as possible.

Decisions about compensation, however, are only part of a larger employment contract or relationship. Such a contract involves an agreement by an employee to perform certain tasks in return for current and future pay from the employer. Contracts are often analyzed in economics using the terms principal and agent. Principals are parties that hire others to help them achieve certain objectives, while agents are those hired to advance the interests of others. In the case of the labor market, the firm can be referred to as the principal and the worker as an agent.

What form do employment contracts take? Such contracts are different from the formal contracts individuals sign (e.g., when taking out an auto loan) in that they are typically incomplete and implicit. Employment contracts are incomplete because they do not specify all the actions that may be required of an individual. This gives the firm some flexibility in responding to changing business conditions. Employment contracts are also implicit because they include a number of informal understandings that can be difficult to legally enforce, and each side typically has the flexibility to end the relationship at any time. Given this structure,
employment contracts have to be self-enforcing to be effective; i.e., each side must perceive it to be in their own interest to live up to the contract. A self-enforcing contract can be particularly hard to achieve when information in the market is asymmetric. This occurs when one party knows more about its intentions or performance under the contract than the other. Such asymmetries increase the likelihood of cheating on the employment contract, so many firms try to structure the terms of the employment contract in such a way that workers will signal their true abilities and intentions.

However, even when employers are successful in choosing workers with the right characteristics, there are incentives for opportunistic behavior on both sides of the employment relationship. Thus it is necessary to structure employment contracts such that both sides are induced to keep their promises about behavior, or in other words, so that it is in both parties’ best interests to continue the employment relationship. These incentives are strongest when the employment relationship creates a surplus (the gap between the employee’s marginal revenue product and his alternative wage) that is shared between the worker and the employer, and thus both sides benefit from the relationship. Surpluses generally arise from investment by the firm in training or through the firm’s reputation as a reliable employer (which can then attract higher-quality workers).

What specific employment practices constitute the substance of an employment contract? Most of the key components of the employment contract are designed around the goal of motivating workers to exert their best efforts. Decisions related to hiring standards, supervision, compensation, and fair treatment of workers all play an important role in maximizing worker productivity.

An obvious way to motivate workers is to pay based on performance. Pay schemes based on individual output include piece-rate and commission pay. But whether this is effective depends first on whether the worker sees his performance as directly linked to his own effort. If performance relies in part on other workers, machines, or other factors, the worker may not be induced to optimal effort. Secondly, the output measure must be quantifiable and relate to the employer’s objectives, and that may be difficult, particularly in service-related industries.

The most common method is to pay employees for their time. Such plans, however, may not lead workers to exert their best efforts without close supervision, unless workers are motivated by group loyalty or future considerations. Guaranteeing workers a stable wage is thought to decrease work incentives in the same way that insuring individuals against unfavorable outcomes reduces their willingness to take steps to avoid losses. This type of behavior is called the moral hazard problem. Essentially, since the worker is being paid for time rather than effort, why put in effort?

The role of the individual worker as part of a group is also an important issue in motivation. Workers who feel that they are not treated fairly relative to others may quit or reduce effort. Workers who feel loyal to their group may expend additional effort. Compensation methods such as gainsharing, profit sharing, and bonus plans tie at least part of an employee’s pay to group achievements.

Given the variations in earnings that are sure to result from output-based pay, employees tend to prefer time-based compensation. Variations in earnings lead to lower levels of average utility if the individual is risk averse (see Appendix 8A). Holding all else constant, a compensating differential is expected to accompany output-based pay. However, it is possible that piece-rate pay schemes are likely to attract more productive workers, since those who gain the most from these plans are those who are more productive. Thus, workers paid under these plans are likely to earn more than time-rate workers, both because they are more productive and due to the compensating differential.

From the employer’s perspective, it is the time-based compensation that leads to the more variable profit stream as well as increased screening and supervising costs. While these considerations suggest that the firm would prefer output-based pay, such schemes are not without their problems. Individually based incentive pay encourages the employee to work hard with little supervision, but such schemes may cause workers to focus so much on the quantity of output that quality suffers, equipment is damaged or misused,
and group loyalty is diminished. Setting the rate for piece-rate work can also be difficult if it involves
direct observation of workers. Some jobs are also not conducive to individually based incentive pay since
the individual’s output may be difficult to measure objectively. Even when there are clear measures of
performance, tying pay to these measures may cause workers to ignore other aspects of the job.

It is the problem of measuring individual output that ultimately gives rise to pay based on group incentives.
Group incentive-pay plans can take many forms. Some tie the pay of the group to gains in productivity,
product quality or cost reductions achieved by the group. These are commonly called gainsharing plans.
Profit sharing and bonus plans link pay or bonus payments to the firm’s overall profit level. In some cases
workers may even own the firm and so split the profits among themselves. The success of group incentive
schemes depends on the ability of the group to control **shirking**. Particularly in large groups, the success of
the group does not usually hinge on the performance of any one individual. Knowing that, some individuals
may realize that they can reap the same financial rewards while doing less. This is an example of a **free-rider**
problem. Of course, if everyone behaves in this way, shirking (free riding) ultimately undermines the success
of the group.

One way to control shirking may be to try to hire only those workers who signal a predisposition toward
cooperation and group loyalty. Alternatively, a firm may take steps to build group loyalty among existing
workers. When feasible, it may help to keep the group size small so that shirking can be more easily detected
and more quickly punished. However, this can conflict with the need for large team production processes.

Profit-based bonuses are most frequently used in the compensation of top executives. Such performance
attainment plans usually tie the executive’s bonus to the profit performance over a 3- to 5-year period.
Basing the award on performance over a shorter period might create incentives to advance the short-run
success of the firm at the expense of long-run profits. Longer periods are thought to involve such a long
wait for compensation that the executive’s performance incentives are reduced.

In an effort to keep executives focused on long-term consequences, firms increasingly tie compensation to
the value of the company’s stock. In general, it is more difficult to fool the stock market into thinking
the earnings prospects of the firm have really been increased through strategies that only serve to increase
short-run profits. The evidence suggests that such compensation plans do help to improve the performance
and stock market value of the firms that adopt them. Attempts to tie the yearly pay of executives to changes
in the stock market value of the firm are typically constrained by the risk aversion of the executives. Given
the volatility stock prices can display in the short run for reasons beyond an executive’s control, firms
that use such schemes may have trouble attracting new executives without a compensating differential.
Additionally, some recent evidence suggests that close relationships between CEOs and their boards of
directors can lead to negotiation of excessive compensation packages, and thus that “outside” directors or
other well-informed agents are necessary to align incentives efficiently.

While incentive pay has the potential for increasing productivity, the litany of problems associated with
such compensation policies has led most firms to continue to rely on time-based pay for much of their
work force. To improve the incentives associated with time-based pay, some firms have instituted systems
involving **merit pay**. These plans attempt to award higher pay raises each year to those workers judged to
be more productive.

Merit-pay plans actually share most of the same problems as output-based pay. For example, how will
performance be measured? Assuming it can be measured, what about factors that are beyond the workers
control? To get around this latter problem, many merit-pay systems rely on the relative ranking of workers
by supervisors. Such a ranking however, can create incentives for workers to be uncooperative or disruptive
towards coworkers, or divert effort into attempts to win favor with supervisors. Because such plans rely on
the subjectivity of managers, they are often distrusted by workers, particularly when evidence surfaces that
managers are being inconsistent in their rankings. To avoid controversy, managers may rely too heavily on
the middle categories of the system, thus mitigating the increase in work incentives that the plan is
designed to foster.
Given the difficulties inherent in whatever method is used as a basis for pay, it is clear that there is no system that will work for all firms. Rather, each firm must determine the appropriate basis after weighing all the costs and benefits. But the basis of pay is not the only decision needed to form a compensation policy. The level of pay is also an important element of the policy, and it too can have an important effect on worker productivity.

The level of pay is often a question of how to divide the surplus created when a worker and firm agree to enter an employment contract. An example of this occurred in Chapter 5 in the context of firm-specific training investments. The surplus consisted of the difference between the worker’s marginal product at the current employer and the real wage a worker could expect elsewhere. The task of the firm was to set the wage so that it could recover the training investment given the expected job tenure of the worker while still keeping the wage above the worker’s alternative offers. But what if the size of the surplus is a function of the wage that is set? That is, what if the worker’s marginal product increases with the wage paid to the worker?

There are good reasons for suspecting that such a relationship between the wage and the marginal product exists. Higher wages expand the pool of applicants and allow the firm to be more selective in hiring. A larger gap between the wage and the worker’s alternative offers should also help to reduce shirking and quits. A reduced quit rate, in turn, tends to lead to more firm-specific training opportunities. To the extent that higher wages are perceived as fair treatment, they may help to build commitment and group loyalty.

Exactly how much the wage should be raised above the market wage in these instances ultimately depends on the marginal costs and benefits of the change. The level that just equates the marginal benefit of continued increases with the marginal cost is called the efficiency wage. It is important to note that an efficiency wage strategy makes sense only when the firm expects a long-term employment relationship with the workers. For this reason, efficiency wage strategies are thought to be more likely in situations where firms utilize an internal labor market strategy (see Chapter 5).

Firms that foster long-term employment relationships through internal labor markets also have some discretion over the sequencing of compensation. Can this discretion be used to increase worker productivity? One productivity-enhancing technique may be to pay workers less (in real terms) than their marginal product early in their careers with the firm, and more later. Such a deferred payment scheme should increase work incentives early on since workers will not want to risk getting fired before they become eligible for the overpayment. Such a scheme would also be most attractive to those workers who are inherently very diligent workers and feel that the risk of being fired is very low.

This scheme will be acceptable to workers and the firm provided the present value of the real earnings stream is equal to the present value of the workers’ marginal product over time. This condition, in turn, requires that for any firm using a deferred payment scheme, the present value of the overpayment must equal the present value of the underpayment. It must also be true that the present value of the earnings stream is greater than or equal to the present value of alternative streams that workers could have in other jobs (or they will leave), and that the firm must earn a normal profit. For profits less than zero, the firm will eventually exit, while supernormal profits attract entry, and thus neither state can be an equilibrium.

Example

Consider a firm where the marginal product of the typical worker varies over time according to the schedule in Table 11-1 where year 0 represents the current year. In order to motivate its workers to exert their best efforts, suppose this firm plans to sequence the pay in such a way that workers receive less than their marginal product early in their career, and then more than their marginal product toward the end of their career. If the firm has decided to pay workers the real wage values shown in Table 11-1 for years 0, 1, and 2, what is the minimum the firm could pay workers in period 3 and still have the plan be acceptable to the workers? Assume the real interest rate is 6%.
Table 11-1

<table>
<thead>
<tr>
<th>Year</th>
<th>Marginal Product</th>
<th>Real Wage</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>10</td>
<td>6</td>
</tr>
<tr>
<td>1</td>
<td>15</td>
<td>12</td>
</tr>
<tr>
<td>2</td>
<td>18</td>
<td>18</td>
</tr>
<tr>
<td>3</td>
<td>20</td>
<td>?</td>
</tr>
</tbody>
</table>

Setting the present value of the underpayment equal to the present value of the overpayment yields:

\[ 4 + \frac{3}{1 + 0.06} = \frac{W_3 - 20}{(1 + 0.06)^3} \]

\[ \Rightarrow W_3 = 6.83 \]

\[ \Rightarrow W_3 = (6.83)(1.06)^3 + 20 \Rightarrow W_3 = 28.13. \]

The schedule of wage and marginal product values over time is shown in Figure 11-1. These schedules each yield identical present values of 56.96 even though the marginal product values only sum to 63, while the wage values sum to 64.13. The actual amount of the underpayment (7; area A + area B) must be less than the actual amount of the overpayment (8.13; area C) to reflect the time value of money.

Figure 11-1

What would happen, however, if the worker were laid off before reaching period 3? The result would be that the worker would never recover the earlier underpayment, and so the present value of the marginal product stream (what the worker could get elsewhere) would exceed the present value of the compensation received. One way this can be avoided is for the employment contract to guarantee that those with the shortest job tenures will be the first to be laid off if the firm cuts back its employment. Providing some of the overpayment in the form of vested pension benefits also could help to guarantee that workers receive what is coming to them.

Alternatively, what would happen if the worker wanted to stay with the firm after period 3 and the firm could not refuse or lower the wage? In this case, the present value of the overpayment would exceed the present value of the underpayment and the firm would not be able to earn even a normal profit. Given that mandatory retirement for most workers is currently illegal under the Age Discrimination in Employment Act, firms have little recourse when older workers want to stay with the firm. Most firms respond to this problem by providing financial incentives for earlier retirement or structuring defined-benefit pensions so that the present value of lifetime pension benefits declines as retirement is postponed past a certain age.
If successful, a deferred payment scheme should eventually lead to a higher present value of compensation than can be attained elsewhere. The workers’ marginal product values at the firm should be shifting up over time if the workers are motivated to exert extra effort by the timing of compensation.

Firms can also affect the timing of compensation through the use of **promotion tournaments**. In such a tournament, employees compete against one another over a number of years for a prized position involving high salary and prestige. There can be only one winner. Such a scheme is similar to a deferred payment in that employees are induced to work hard now by the expectation of a large bonus later. The difference is that not everyone will actually receive the bonus. Therefore, such a scheme will be acceptable to workers only if the losers are also treated relatively well, i.e., their wage is allowed to exceed their marginal product. This helps to explain why some firms continue to employ a few workers that are easily identifiable as relatively unprofitable employees.

Tournaments create incentives for effort depending on the marginal benefits that workers expect to receive from additional effort. Marginal benefits depend on the size of the prize and the increase in the probability of winning that comes from more effort. Thus where additional effort is very costly or where extra effort is perceived to have little effect on winning the prize, prizes must be very large to create incentives. Like piece rates, tournaments may have value to the firm because they can cause more productive or more motivated workers to sort themselves into such jobs.

However, tournaments can also result in efforts to sabotage others rather than to improve one’s own chances, and thus may not be appropriate where team effort is important. Also, after the tournament is over, workers may have no further incentive for effort.

The willingness of employees to work very hard early in their career is enhanced when workers view themselves more as a member of a profession or field as opposed to a member of a particular firm. In such situations, hard work may not only impress the current employer but may also improve employment prospects at other employers. When workers are highly motivated by such **career concerns**, firms may not need to employ incentive pay or pay sequencing efforts to enhance productivity.

The last part of Chapter 11 tackles two empirical puzzles using the theories of pay and productivity discussed earlier in the chapter, along with theories from previous chapters. The first puzzle is: Why do earnings increase with a worker’s job tenure? Earlier chapters suggest that longer tenure and higher wages both reflect more productive matches between worker and employer. A second reason involves firm-specific investments, jointly undertaken by both workers and employers. The discussion of pay sequencing in this chapter suggests that another explanation is that higher wages earned by those with longer job tenures serve as a reward for enduring underpayment earlier in their careers.

The second empirical puzzle is: Why do large firms pay higher wages? The explanation contributed by Chapter 11 is that the complicated and interdependent production processes at most large firms require the payment of efficiency wages to reduce shirking. Also, large firms are better positioned to use internal labor markets given the wide variety of positions available in a large firm. Internal labor markets, in turn, increase the likelihood of efficiency wage and pay sequencing strategies. These strategies raise the marginal productivity of workers above what it would be in other firms, thus increasing the overall level of pay. Such productivity enhancing strategies should appeal to large firms in that it is costly to supervise a large workforce.