

Professor Allen
EC 507
9/1/2005

- I. Production
- II. Principal-agent problem
- III. Factors influencing effectiveness of pay incentives
- IV. Economics of compensation
 - A. Competitive factors: human capital, compensating differentials
 - B. Efficiency wages
 - C. Internal labor markets
 - D. Individual incentives
 - E. Group incentives
 - F. Profit sharing
 - G. Options

I. Production

A. Production function

Goal here is to begin understanding determinants of costs. Firms spend money on land, labor, material, equipment, and structures and then use these inputs to produce goods and services to sell in the marketplace. To effectively manage costs, firms must (1) select the right combination of inputs, (2) pay the right price for key inputs and (3) employ those inputs efficiently.

Before looking at these decisions, we need to be able to describe the technology, which to an economist means the relationship between inputs and output. (To keep things simple, we will assume that the underlying technology is stable; we will alter this assumption later on.) This relationship is called a *production function*. Letting Q =output, L =labor, and K =capital, this relationship is usually written as

$$Q = F(K,L).$$

The production function describes the maximum level of output that can be produced with given amounts of capital and labor. For instance in investment banking, Q might be the dollar value of leveraged buyouts, mergers, and acquisitions; L would be hours worked by MBAs, lawyers, and accountants; K would be rental value of computers, phone systems, and office space. In education, Q might be degrees granted or students taught; L would be hours worked by teachers, administrators, and staff; K would be rental value of classrooms, labs, and school buses. (There is a certain resemblance between a production function and a recipe – take so many chocolate chips, so much flour, and so much butter and you get so many chocolate chip cookies.)

The production function also describes tradeoffs between K and L that produce the same Q . Nike can produce shoes with expensive, high-tech equipment and relatively little labor or with less sophisticated equipment and lots of handwork. Highways and other public works can be built with expensive earth-moving and paving equipment and small work crews or with hand tools, forms and massive work crews. You can see these tradeoffs by comparing old vs. new facilities in the US in the same industry or by

comparing production methods in the US, Japan, and W. Europe to those in Asia and Latin America.

In practice the production function can include other types of inputs (e.g., land, materials, R&D), can have multiple types of K (e.g., equipment vs. structures, high-tech equipment vs. other equipment) and can have multiple types of L (e.g., different skill levels as indicated by measures such as occupation, education, experience, etc.). A key research issue in economics and management today is the impact of computing capital on production. To examine this, researchers will want to measure both computing K and non-computing K, as well as be able to control for other determinants of productivity that are potentially correlated with computing K (e.g., skill of L). Regression analysis is used to estimate production functions.

A few more definitions:

- 1) Short run vs. long run: In the short run some inputs are fixed. In the long run all inputs are variable. Most organizations are locked into their capital and land decisions for some time frame. The length of this time frame will vary across firms and industries. Some firms consciously decide to rent or lease everything possible so that they have maximum flexibility to move rapidly to new market opportunities. Others consciously decide to make long-term commitments to specific types of equipment or to specific suppliers of materials. In construction, the long run could be as short as the length of the lease or rental on equipment whereas in electricity generation, it could be decades. (Aside: in some unionized companies in the US and in most companies in W. Europe labor is a fixed cost because firms are effectively prohibited from firing workers.)
- 2) Average product: This is simply the ratio of Q to one of the inputs. We commonly refer to the average product of labor (AP_L), which is simply the ratio of output to labor input, or

$$AP_L = Q/L$$

This term is also known as *labor productivity*. This is a measure that is repeatedly misused in most discussions about economics. Labor productivity depends on all inputs, not just labor – in fact the surest way to get high values of labor productivity is to adopt a technology that is very capital-intensive. To track productivity changes over time, economists use a measure called *multifactor productivity*, which has a weighted average of L and K in the denominator.

$$MFP = Q/(s_L L + (1 - s_L)K),$$

Where s_L = the ratio of labor cost to total cost. This measure controls for the level of K, the level of L, and the ratio of K/L – so that changes in MFP can be interpreted as true changes in productivity. This is an important issue because ultimately the change in our overall standard of living depends on the overall measure of MFP for our society. Also firms that are seriously committed to measurement will want to use MFP instead of AP.

- 3) Marginal product: This is the change in Q that results from a one unit change in one of the inputs, holding the other inputs constant. The marginal product of labor (MP_L) is defined as

$$MP_L = \Delta Q/\Delta L$$

Aside: For those who remember their calculus of two or more variables, a more precise definition of MP_L is that it is the partial derivative of Q with respect to L, holding K constant, or

$$MP_L = \delta Q / \delta L$$

- 4) Returns to scale: This is what happens to Q when you change K and L by the same proportion. For instance if Q doubles when you double K and L, you have constant returns to scale. If Q more than doubles, you have increasing returns to scale. If Q less than doubles, you have decreasing returns to scale. This is obviously an important concept for managers. Constant returns are easy to understand – one dry cleaning operation can easily replicate itself in every neighborhood. Increasing returns may result from greater specialization of labor or capital at large scale (e.g., more sophisticated technology) or fundamental engineering relationships (e.g., double a pipe's circumference and you quadruple throughput). Decreasing returns often result from coordination and communication difficulties.

B. The relationship between output, AP_L , and MP_L .

Figure 6.1 in the text (p. 192) describes a typical production function with one variable input (L). We can measure AP_L by using the slope of a ray (dotted lines) going from any point on the production function to the origin (Why? Slope of dotted lines is ratio of Y over X, which is Q over L, which is AP_L) and MP_L by using the slope of the production function itself. The function is first convex with respect to the origin. In this region both AP_L and MP_L are increasing as the firm is able to take advantage of specialization in labor (classic example: Adam Smith's pin factory). At point B MP_L reaches its maximum level. At point C the function becomes concave and AP_L reaches its maximum level at C.

The key economic concept that comes out of this two panel diagram is a very simple one: that eventually MP_L must decline as more units of L are added. This is called the *law of diminishing returns*. It is a law because we are holding some inputs fixed (let's call them K) and as more and more units of labor are added, the ratio of K to L declines. The fact that each worker has less K to work with can be offset at first by taking advantage of specialization and improved organization, but somewhere these economies must be overridden by the simple arithmetic of less K for each L. This means that eventually MP_L must decline as output expands.

C. Making choices about inputs

1. Tradeoffs between K and L: In equilibrium, firms will make decisions about inputs by comparing their relative productivity to their relative cost. For instance if

$$MP_K / MP_L > c/w \quad (\text{or } MP_K / c > MP_L / w),$$

where c =cost of capital and w = cost of labor, then the firm needs to add more K and use less L. Why? One way to see this is that the productivity advantage of using more K is greater than the cost disadvantage. Another is to see that K yields more revenue per dollar spent than L. Similarly, if

$$MP_K / MP_L < c/w \quad (\text{or } MP_K / c < MP_L / w),$$

Then the firm needs to add more L and use less K. The firm will have the right balance between L and K when

$$MP_K / MP_L = c/w \quad (\text{or } MP_K / c = MP_L / w).$$

2. Choices about different types of L: A firm might be wondering whether to require all new hires in its sales groups to have college degrees. College grads would have to be paid higher salaries, so they would have to be more effective salespersons to compensate. The decision rule for the firm:

Switch to college grads if $MP_C / MP_{HS} > w_C / w_{HS}$;

Stick with high school grads if $MP_C / MP_{HS} < w_C / w_{HS}$.

3. Choices about plant location: Should Nike make athletic footwear in Oregon or Indonesia? Presumably labor is cheaper in Indonesia but more productive in Oregon. The key is to compare the productivity advantage of Oregon labor to the cost advantage. The decision rule:

Indonesia if $MP_O/MP_I < w_O/w_I$

Oregon if $MP_O/MP_I > w_O/w_I$

II. Principal-agent problem

The view of the world embodied in the production function seems to imply that management practices and policies have no influence on productivity. In this view of the world, production functions are like recipes – so many units of K plus so many units of L and out pop so many units of output. As any experienced cook knows, this is all fine in theory, but in practice you better watch what is going on in the oven to make sure nothing gets burned. More importantly, you want to make sure that the incentives of the cook are aligned with the overall goals of the enterprise. If a baker gets paid by the cookie, you are likely to see the world's smallest cookies ever. More fundamentally, the organization needs to make sure that the baker has the proper incentives to make sure that customers are satisfied and cookie sales are profitable. In this section of the class we will show (1) that the interests of top management and employees are not always aligned with the interests of the owners of the firm (who in most cases are shareholders) and (2) incentive schemes can be used to bring these interests in closer alignment.

There are three forms of ownership for private enterprises: proprietorship, where a single person owns the firm; partnership, where two or more individuals own the firm; and corporations, where the firm is owned by shareholders. (We will not spend any time on the difference between partnerships and proprietorships except to note that the only important difference is the potential for disagreement among the owners of a partnership.) One important difference between the first two forms of ownership and the last is liability. If business turns bad or the business loses a court case, partners and proprietors stand to lose not only their business but also their own personal assets. In a corporation, the worst thing that can happen is the stock price goes to zero; stockholders will not have to hand their car keys over to the repo man (unless they cannot pay their margin calls).

Typically in partnerships and proprietorships, the owner(s) manage the firm. This means that they have every incentive to minimize costs because they get to keep the profits. In corporations, the CEO is an employee and her responsibility is to manage the firm in the interests of the shareholders. There are a number of reasons to think that even if the CEO were paid a competitive salary, she would want to have the shareholder's best interest at heart:

- The carrot: There is a competitive labor market for CEOs. If she does a good job meeting the stockholder's needs, she will start getting job offers from other corporations and can either accept the offers or use them as leverage to raise her salary.
- The first stick: The CEO reports to the board of directors, who can fire the CEO if her performance is judged inadequate.

- The second stick: Even if the CEO and a majority of the board all went to college together at Yale, the entire company is still subject to the discipline of the market. A corporate raider could buy the firm and replace management.

Nonetheless, how can the board of directors (much less the average shareholder) know that the CEO is doing everything she can for the bottom line? CEOs could easily want some things that run counter to profitability, including large staffs, lavish offices, and acquisitions that get the CEO lots of media buzz but subtract from the bottom line, e.g. Carly Fiorina's decision to merge H-P with Compaq. CEOs can engage in such behavior because no one – the board, financial analysts, corporate raiders -- has the resources to monitor all facets of their behavior. Knowing that, CEOs have some flexibility to pursue their own agenda.

Recently some extreme cases of CEOs in handcuffs have come to light. In these cases, CEOs have gone beyond pushing the envelope of the principal-agent problem – they have broken federal laws. You don't have to go to such extremes as using company funds to throw a lavish birthday party for your wife on a Mediterranean island (Tyco) or reporting phony quarterly and annual financials (Worldcom) to take advantage of your stockholders. You can do it perfectly legally by spending corporate resources on things that add value to the CEO's ego but do not enhance the value of the firm.

I have focused so far on the CEO, but the same argument applies to any employee. You can tell your boss that you are meeting a new client between 1 and 3; how will he know that you are checking out the sale at Belks or having a long lunch at the Cheesecake Factory? A purchasing agent can steer business to suppliers who provide the proper "grease," e.g., a case of Grey Goose come holiday time.

These are all examples of what economists call the *principal-agent problem*. The agent is the person who acts on behalf of the principal, whereas it is the principal whose welfare is affected by the action. In most cases the principal has ownership or property rights. Here are some examples of principal-agent relationships; be sure you can explain in each case the problems that might arise:

Principal	Agent
Shareholders	CEO, employees
Building owner	Building contractor
Stock owner	Stockbroker
Landowner	Sharecropper

The essence of the principal-agent problem is that agents will try to pursue their own interests, at the expense of the welfare of the principal. These problems arise for three main reasons:

- 1) **Observability:** the owner lacks the time and other resources to observe all aspects of the agent's behavior. How does Madonna know she has gotten the best possible deal when her agent could have settled on a lower contract in return for getting two other acts he represents signed with the same record label?
- 2) **Moral hazard:** knowing that not all actions will be observed, the agent can *pursue his own agenda*. We have already discussed this in terms of CEO behavior, but it can apply to a much wider range of employees and a much broader set of situations.

Anyone with spending or hiring authority in a firm has the potential to put individual interests above those of the stockholders.

- 3) Asymmetric information: An additional factor that sometimes comes into play is that the agent may have strategic information that the principal does not have. R&D employees know which projects are the best prospects and which are not so great, but they tend to present all options to management with great enthusiasm.

(More examples: read chapter 2 of the current bestseller *Freakonomics* by University of Chicago economist Steve Levitt. These same concepts apply to high school teachers subject to high stakes testing, sumo wrestlers, and office workers.)

At first glance there would seem to be an easy solution to the principal-agent problem for a firm and its employees: change the pay system from straight salary to one in which the CEO and the employees have their fortunes tied to the value of their own performance. This would certainly align incentives of employees with the interests of shareholders. But there are two big catches:

- 1) Employee performance is not accurately measured. Objective measures are not available for many types of work; subjective measures are subject to manipulation. This is a huge topic that we will not go into here; it is examined in BUS 530 and ACC 580.
- 2) Employee performance is partially determined by factors that are not subject to the employee's control, e.g., Carolina Panther's coach John Fox had a winning game plan against the Eagles two years ago in the regular season (before they beat them in the playoffs on the way to the Super Bowl!) but one reason the team lost is that normally reliable John Kasay missed three field goals and an extra point.

The basic point is that the dollar value of any person's contribution is subject to random shocks. There could be breakdowns in equipment, bad weather, consumer demand could fall – any of these factors could cause the dollar value of an employee's contribution to drop precipitously. Performance depends on effort AND random shocks. *So pay for performance really means pay based on both effort and luck.* Under any pay-for-performance scheme, employees are going to be exposed to a much greater variation in income than they would have been if they had been paid a straight salary. In essence, performance-based pay shifts some risks away from the firm toward the employee.

Assuming that most employees are *risk-averse*, this is not a very efficient arrangement for risk management. Remember that shareholders face limited liability and can purchase diversified portfolios of stocks to manage their risk. Extreme forms of performance-based pay put employees in a very poor position in terms of risk exposure. Consider employee stock ownership plans. Most employees have a very sizable investment in human capital that is specific to the firm (training and skills they have developed that are not valuable anywhere else). If their pay becomes heavily contingent on the financial performance of the firm, they are subject to a double whammy if the firm suffers an extremely adverse shock – they lose their jobs and their financial holdings plummet in value. Prime recent examples: Enron employees, honest Arthur Andersen partners.

Bottom line: in designing incentive schemes to motivate CEOs and employees, there is a tradeoff between solving the agency problem and providing an efficient level of

insurance against adverse incomes. We now turn to exploring this tradeoff in more detail.

III. Factors influencing effectiveness of pay incentives

One way to solve the principal-agent problem is to make the agent's pay contingent on some observable measure of performance. In practice this can mean some form of incentive pay, including bonuses, commissions, and stock-based compensation (including options). We will refer to these collectively as high-powered performance incentives. The following table is slightly modified from David Besanko, David Dranove, and Mark Shanley's book *The Economics of Strategy*:

<i>Factors favoring high-powered performance incentives</i>	<i>Factors favoring low-powered performance incentives</i>
Employees have low level of risk aversion	Employees have high level of risk aversion
Low probability of random shocks gives employees control over performance	High probability of random shocks gives employees little control over performance
Employees have low level of effort aversion	Employees have high level of effort aversion
Effort is costly to observe	Effort is easy to observe
The marginal contribution of effort to performance is high	The marginal contribution of effort to performance is low
Relatively noise-free performance measures	Relatively noisy performance measures

Risk-aversion: as we will discuss in more detail below, with a competitive labor market, employees demand extra payment for jobs with undesirable working conditions. An extremely risk-averse employee will demand a much higher performance premium than one who has little risk-aversion. Accordingly, firms that offer high-powered incentives try to attract employees with relatively low levels of risk aversion.

Probability of random shocks: holding one's own preference toward risk taking constant, the greater the risk → the greater the premium. Think of it this way. Suppose you have job offers from two firms – a telecom that has had its ups and downs and Microsoft. Each offers you the same salary and will pay you a bonus of 200% of your salary if you meet specified performance targets. Would you have a hard time making up your mind between these two offers?

Effort aversion: some folks are extremely averse to extra effort, so they would require especially high premiums to get them to put forth that effort. This is an important point to remember if you are trying to install high-powered incentives in an organization where the culture is extremely laid back.

Cost of observing effort: remember that lack of observability is a key source of the agency problem. If you can monitor your employees at relatively low cost, then you do not need high-powered incentives to motivate them. Instead you can correct their

behavior and dismiss them if they fail to shape up (and there are replacements available).

Impact of effort on performance: in some jobs extra effort and the extra performance generated by that effort simply do not have that much impact on the bottom line. The housekeeping staff at the Red Roof Inn must meet an acceptable level of performance in cleaning the rooms, but customers are not going to pay a premium for ultra-shiny doorknobs. So why should Red Roof pay an effort premium? (There might be another reason; see below.)

Noise in performance measures: if the measures amplify the variation in observed performance (and in pay as well), then an additional risk premium will be required. This extra cost makes high-powered incentives less desirable in this situation.

Looking across all of these factors, we can see why firms make heavy use of stock grants and options to compensate their CEOs. They want individuals who are willing to take prudent risks and work extremely long hours. The impact of their effort on the firm's performance is quite high. Since their charge is to oversee wealth creation for shareholders, stock value is an appropriate performance measure. It is also easy to see why there might be an issue in trying the same approach with middle managers, much less the crew on the loading dock.

IV. Economics of compensation

We now look at some specific compensation schemes and how they deal with the agency problem. In each case be sure to think about how the scheme affects

- Behavior of incumbent employees (incentive effects)
- Characteristics of employees seeking to join your organization (sorting effects)
- Risk allocation between the shareholders and the employee (risk-sharing effects)

A. Competitive factors: human capital, compensating differentials

A quick primer on labor economics: pay rates depend on supply and demand. If all workers were alike they would get paid the same. Workers who invest heavily in their own education and training receive higher wages than those who invest relatively little. This reflects their greater productivity as well as a payback on their investment in *human capital*. Job characteristics also affect pay. Employers providing undesirable working conditions (e.g., physical dangers, extreme stress, lots of travel) must pay more than other employers to attract workers. This pay premium is called a *compensating wage differential*; this idea goes back to Adam Smith's *Wealth of Nations*, 1776.

Incentive and sorting effects: not really applicable, pay based on skill and working conditions is necessary to attract qualified workforce

Risk-sharing effects: none

B. Efficiency wages

Let's return to the simple competitive model and look at a set of workers with identical amounts of human capital and employers who provide identical working conditions. Then all of these workers should get paid the same amount (w^*) and would presumably have the ability to leave one employer for another without any

problem. So why should they work very hard? The worst that could happen is they get fired and get another job at exactly the same salary!

One way for employers who value effort to stop workers from shirking would be to pay them $w > w^*$. Then if they got caught shirking and were fired, they would stand to lose $w - w^*$. As long as the wage premium ($w - w^*$) is greater than the cost of effort, then these employers would not have to worry about shirking any more. The wage at which shirking disappears is called the *efficiency wage*. The increase in productivity pays for the higher costs of labor. (There are two ways this could work: you could have a dual labor market where some firms pay the wage premium and others pay w^* and there would be no unemployment OR all firms could pay the wage premium and there would be fewer jobs so fired workers would be unemployed – the text presents the latter case.)

In practice wage premiums raise productivity through a variety of mechanisms besides greater effort. One way is through reducing turnover, which allows the firm to invest more heavily in training and by reducing the costs of finding and replacing exiting employees. Another way would be through discouraging workers from joining unions. Yet another would be attracting highly motivated workers.

Incentive effect: more \$ for more work

Sorting effect: attract better workers

Risk-sharing effect: not applicable

Classic example: Henry Ford's \$5 day in 1914

C. Internal labor markets

Despite all we hear about job hopping and the death of loyalty, a surprisingly large number of people work for the majority of their career in a single organization (see my paper on job security in large corporations). In corporations such as UPS, we observe the following:

- Many long-term employees
- Hiring focused on entry level
- Promote from within
- Training encouraged through job progression
- Salary based mainly on what job you have
- Seniority matters a lot
- Due process procedure for resolving disputes (Teamsters grievance procedure for union members; internal procedure for all others)

This sounds so retro! Granted in this model, firms can end up becoming overly protective of low performers and have difficulty keeping potential stars. There also is the risk of developing an inbred culture that is resistant to change. But there are advantages including lower turnover costs, more incentives to invest in training and development, better teamwork, better information about internal than outside applicants, and greater ability of internal applicants to hit the job running.

How does the pay scheme work in this type of setting? So far we have assumed that wages must reflect productivity in the current period because at the end of each period, either side may recontract. Consider the following twist: the firm pays the worker less than his output in the early stages of his career and then pays more than

his output in the later stages. If the worker sticks around for the entire career, the two even out.

But now the worker realizes there is a tremendous incentive to stick around for his entire career because then he can get paid more than he produces (and more than he can get paid anywhere else). This makes him less likely to shirk or quit. One catch: the firm needs to have the retirement plan structured so that the worker wants to leave in the later stages of his career. Otherwise the worker could end up getting the better of the bargain.

Incentive effects: encourages training and discourages turnovers

Sorting effects: attracts many but not a turn on for stars

Risk-sharing effects: not applicable

D. Individual incentives

Piece rates and commissions give workers a direct reward tied to their productivity. These schemes usually consist of a guaranteed base wage or salary, plus an extra payment if they produce or sell more than a given amount. In addition to the incentives for effort, this form of pay attracts hard workers who are willing to take risks.

Yet these pay methods are not used for the great majority of employees. One obvious problem is that individual output is hard to objectively measure outside of a handful of occupations – ag labor, sales, apparel manufacturing. Another problem is endemic to all incentive schemes – people do what the employer pays them to do, which is not always the same as what the employer wants them to do. You have to worry about quality (lots of keystrokes may not mean lots of work getting done) and moral hazard (Domino's 30-minute delivery pledge, Sears decision to pay auto center managers a bonus directly linked to total spending by consumers).

Incentive effects: work harder, do what you are paid to do

Sorting effects: more attractive to hard workers

Risk-sharing effects: workers more at risk, shareholders less at risk (but probably not by much)

E. Group incentives

Pay can be tailored to group performance, where the group could be as small as a team or as large as a plant or division. This can be done in a variety of ways, such as paying a bonus if sales, cost saving, safety, quality, or production targets are met. The advantage of this approach over individual incentives is that (1) it is easier to measure group output than individual output in most cases, (2) it encourages teamwork and cooperation, and (3) peer pressure can be substituted for direct supervision. The downside is what I call the *1/n problem*. Each employee gets a share of $1/n$ of the group's reward. If n is a relatively small number, then each person feels a stake in his or her own behavior and the behavior of those around them. But at some value of n , people start to tune out – they see no connection between what they do and their reward. So they become *free-riders* on the efforts of others.

Incentive effects: work harder, more teamwork vs. $1/n$ problem

Sorting effects: attractive to both hard workers and free riders

Risk-sharing effects: workers more at risk, shareholders less at risk

F. Profit sharing

This is a particular type of group incentive where the payment is directly based on the firm's profits. There are a variety of ways to implement this. One would be a cash bonus that gets paid if the firm meets certain profitability targets. Another is a direct grant of stock to employees; this is often used to get people to work at startups for long hours and low pay. The advantage of this approach is that it aligns the interests of shareholders and employees. This has direct financial value and it also may have an important symbolic value.

The downsides (in addition to the $1/n$ problem) are that profits reflect so many things besides worker effort that even in organizations with small n , workers may conclude there is little connection between reward and effort. Also, there is a new observability problem: workers have to trust the firm's top management (and its accountants) to produce an honest measure of profits. Finally, keep in mind the arguments about efficient risk-sharing made above.

G. Options

I used to need five minutes to explain these, but I think it is safe to say that everyone understands how these work now. The theory is similar to profit sharing, but lots of issues come up in terms of implementation, including how long a time horizon for vesting the option, should the option be indexed to some overall measure of stock market or industry performance, and what to do when options are underwater (price of stock is below the strike price).

Incentive effects: work harder, best scheme for aligning worker interests with shareholders as long as worker feels options are in the money

Sorting effects: ask Cisco recruiters from a couple of years ago

Risk-sharing effects: ask Cisco employees now