Economics 463 lecture 18 2005 R.J. Butler
"Compensation: Work Incentives and Implicit Contracts"

I. Introduction

In chapter 12 of Borjas, there are two underlying ideas in discussing work incentives and labor "contracts": 1) that there is asymmetric information that affects the type of economic institutions that we observe (what institutions emerge to sort types of compensation, types of jobs, and types of workers into their most "productive" uses), and 2) there is often some firm specific capital in some workers that makes optimal "implicit" contracts.

With respect to compensation, asymmetric information and implicit contracts affect three dimensions of how individuals are paid: 1) whether on a time basis or piece rate basis (or combination, or even on a gainsharing basis in which all members of a group are paid on the group piece-rate basis) basis, 2) how much to pay (relative to both their market value, and their value to the firm), and 3) how to sequence the pay (whether to delay some pay in the form of higher pensions in the future to generate more employment stability). We address the first two issues with a simple model in section II. Then examine the third issue of the sequencing of pay using two models in sections III and IV: tournaments model, and a delayed compensation (benefits) model. In the lecture 19, we look at efficiency wages and the implicit contract model in some more delay—with a model in which layoffs are possible.

II. One sector model, with two technologies available to produce the homogeneous output (one a piece-rate technology, and the other a time-rate technology), with workers varying in their ability to produce the product on a piece rate basis.

Piece-rate=pay on the basis of number of things produced
time-rate=pay on the basis of number of hours worked

A. The firm: to use a piece-rate or a time-rate technology

Firms differ in their ability to measure (or monitor) output. Workers differ in their ability to produce output; the workers know their individual ability, the firm does not until it hires the workers. Even after hiring the workers, the firm may not fully be able to access or measure the workers output—and hence, their ability--given the technology the firm employs. The firm cannot expect the worker to truthfully indicate his productivity when the firm can't observe it. If the firm can monitor/measure the output of each individual worker at relatively low cost, then they can pay the worker by piece rate (in competitive markets, the piece-rate would equal the marginal product of the worker). If the costs of measuring output is sufficiently high, the firm will pay on a time-rate basis even if it knows workers have some incentive to shirk: as long as monitoring costs under the time-rate system exceeds the loss due to shirking, then the firm will offer time-rate pay. Therefore, piece-rates tend to be paid where monitor costs are low (house repairs, buckets of fruit picked, dollar value of sales made, number of shirts sewn, etc.), and time-rates tend to be paid where monitoring costs are high (teachers, members of software development teams, public administration, etc.).

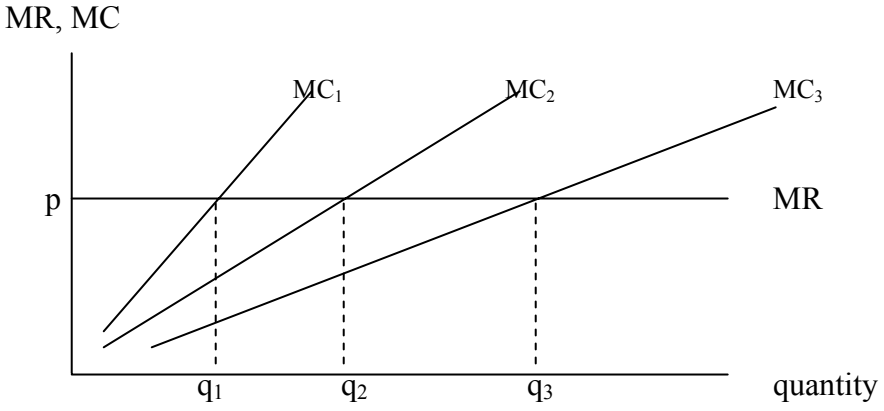
In competitive markets, the time-rate pay will adjust until the wage pays just pays for the value-added generated by the workers on time-pay. This will involve the establishment of a minimum level of productivity that the time-pay worker must achieve, such as showing up to

work and going through a specified lists of activities (even though ones productivity at those activities is difficult to monitor).

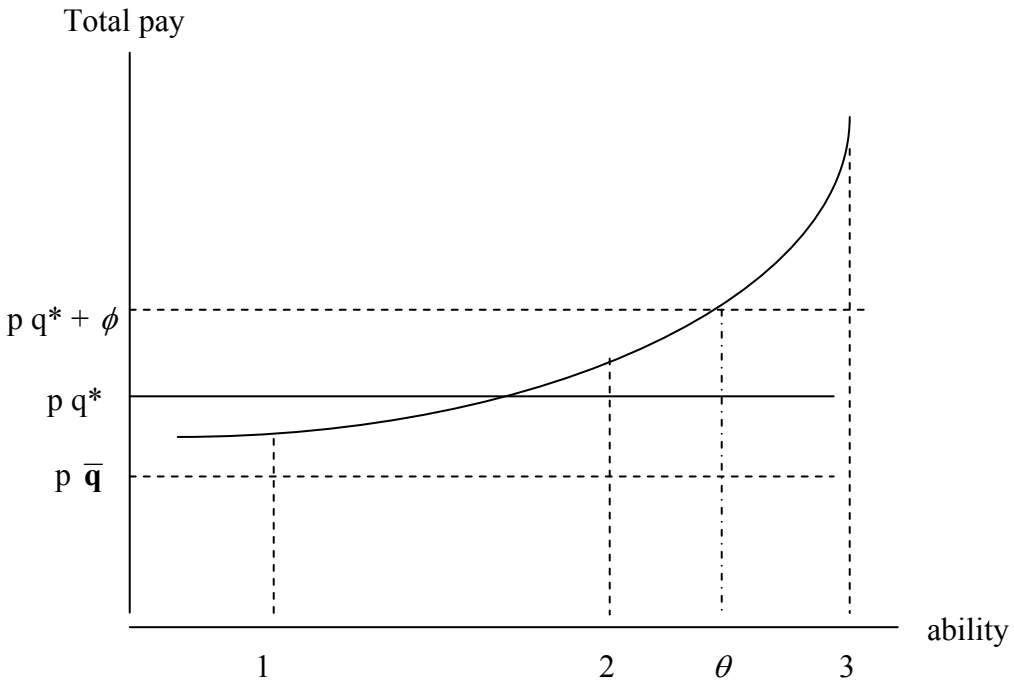
B. The employee: making a choice given risk aversion and ability differences

The time-rate of pay is not contingent upon the number of units produced or sold, and so is the (relatively) risk-free compensation. Piece-rates, on the other hand, depend on the number of units produced and this may vary (for reasons of personal health, differences in aggregate demand--car sales may be slow, etc.). Time-rate pay, on the other hand, insures against these fluctuations in income, and so is favored by risk adverse workers over uncertain incomes with the same expected value. Consider the implications of ability differences in our model, from the workers perspective, in the next two diagrams. Then we return to a brief discussion of the market clearing wage for time-rate workers.

Assume the piece rate given for each unit of the good produced by the worker is p, so this is the marginal revenue to the worker of piece-rate production. Differences between workers arise because of differences in their marginal costs of production: the more ability produce at lower marginal cost. Workers produce at the rate where the MR=MC, as indicated in the diagram below (worker 1 is the least able; worker 3 is the most able with the lowest MC for each and every level of output):



Individual three, the most able, will produce the greatest amount, q_3, per fixed unit of time and will also have the greatest revenue ($p\ q_3$) of the three workers pictured. (Total pay to each individual under a piece-rate compensation scheme equals the piece-rate price times the quantity that each produces.) We assume that the minimal acceptable level of output under the time-rate system is $\bar{\mathbf{q}}$ (this minimal work requirement is assumed to be observable, but omitting this from the model would have no impact on any of the model results). We also assume that pay to those under the time-rate system is competitive in the sense that the time-rate pay yields no rents to the firm, so that the average time-rate pay, p, is just equal to the average productivity of those chosing time-rate work. The compensating variation for uncertain income is assumed to equal ϕ (the same for all workers), so that only those workers whose piece-rate ability exceeds ϕ+p will choose to work for piece-rates, all those with ability less than ϕ+p will be time-rate workers.



This is pictured above: workers 1 and 2 chose to be time-rate workers (the more productive workers like 2 just compensate for the less productive workers 1, on average) and θ is the level of ability that separates the piece rate from the time-rate workers (it provides the cut-off for risk compensating differential that separates the two groups). Worker 3, therefore, is a piece rate worker in this model. A number of results follow from the model:

1. There is a self-selection into the two types of jobs: workers would a greater ability to do piece-rate work tend to choose piece-rate jobs.
2. Piece-rate workers make more money than time-rate workers, other things equal.
3. As workers become more risk adverse, they have a tendency to prefer time-rate pay over piece rate pay (as ϕ goes up , θ goes up).

III. Tournaments

In the piece-rate vs. time-rate model above, workers were paid on the basis of their actual absolute level of their performance or ability (the average absolute performance among the time-rate workers, and the actual performance for each individual piece-rate worker). Tournament theory is used to explain situations in which workers are rewarded on the basis of their relative performance, such as senior executives who vie for the CEO position in a company. In the tournament theory, workers are ranked according to their performance and then rewarded on the basis of the rankings, with the winner receiving a prize (earnings) that is much greater than the

performance difference based on the absolute level of performance. That is, the winner's prize is relatively large given the productivity differences between the first place finisher and the second, third, and fourth place finishers.
Tournament prizes are common in sports (winners of golf tournaments, SuperBaals, prize fights, etc.), but also seem to characterize the race to be a CEO of a company. Such tournaments exist in order to elicit the appropriate amount of effort from the contestants, where their actual performance is more difficult to measure than their relative performance (or rank).

Consider Mary and Eliza who both want to be CEO of BYU, Inc. They are in a tournament to become CEO, and both their own efforts and random events beyond their control will determine the likelihood of becoming the CEO and winning the prize. The winner gets a salary of w_1 and the loser gets a salary w_2, $w_1 > w_2$. Hence the prize spread is $w_1 - w_2$, and the question arises as to how big that spread should be.

Mary's expected utility is assumed to equal her expected earnings minus the disutility of putting effort into participating in the tournament:

$P_M w_1 + (1-P_M)w_2$ - disutility from E_M units of effort,

with a similar expected utility for Eliza. Noting that the probability of winning depends on Mary's efforts, as well as Eliza's efforts, we can write:

$P_M = P_M(E_M, E_E)$

Writing the disutility from effort(for Mary) as Disutility$=G(E_M)$, we can write the expected utility as

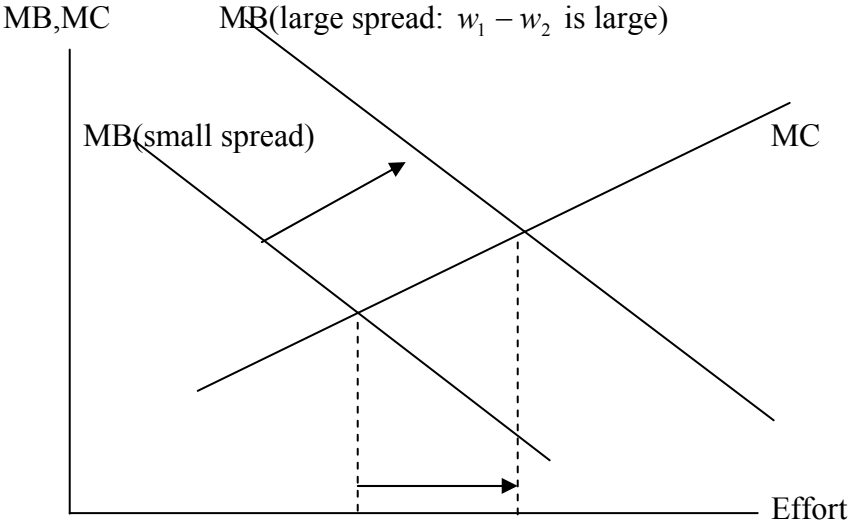
$P_M(E_M, E_E)\ w_1 + (1- P_M(E_M, E_E)\ )\ w_2 \ - G(E_M)$

The first two terms are the expected earnings from being in the CEO tournament; if $P_M=.5$, w_1=\$2million and w_2=\$200,000, then the expected earnings for Mary is \$1.1 million. (Even if she loses she will make as much as a new assistant professor in finance.) Mary's problem is to decide how much effort to give in order to maximize her utility. The solution to Mary's maximization problem is to take the derivative of this last function with respect to EM and set the derivative equal to zero as follows

$$\frac{\partial P_M}{\partial E_M}(w_1 - w_2) - \frac{\partial G}{\partial E_M} = 0 \text{ or} \qquad \frac{\partial P_M}{\partial E_M}(w_1 - w_2) = \frac{\partial G}{\partial E_M}.$$

This last expression indicates the optimal level of effort is such that the marginal gain from additional effort (the gain in the likelihood of winning $\frac{\partial P_M}{\partial E_M}$, given the prize spread, $(w_1 - w_2)$) just equals the marginal cost of the effort, $\frac{\partial G}{\partial E_M}$. As the spread between the first and

second place finish increases, the marginal benefit from effort rises. This will increase the effort spent until the marginal benefit and marginal cost are again equalized. The greater the prize spread, $(w_1 - w_2)$, the greater the effort. This is pictured below:



Eliza faces the same incentives as Mary, and so if they are equally able, will put forth the same level of investment and effort to win the CEO game. The game will likely to be close until the very end. But this is what companies looking to promote individuals to the CEO position want: all the VP and senior VPs in contention need to have the appropriate incentives to work hard until the very end of the game, for the good of the company (else all the contestants in key positions would begin to slack off).

Tournament play can provide too much competition in some settings: in particular, where pay by rank can lead to sabotage of other individuals efforts: on assembly lines or on software teams, pay by rank may be counterproductive as other workers can have a significant affect on your work.

VI. Delayed Compensation

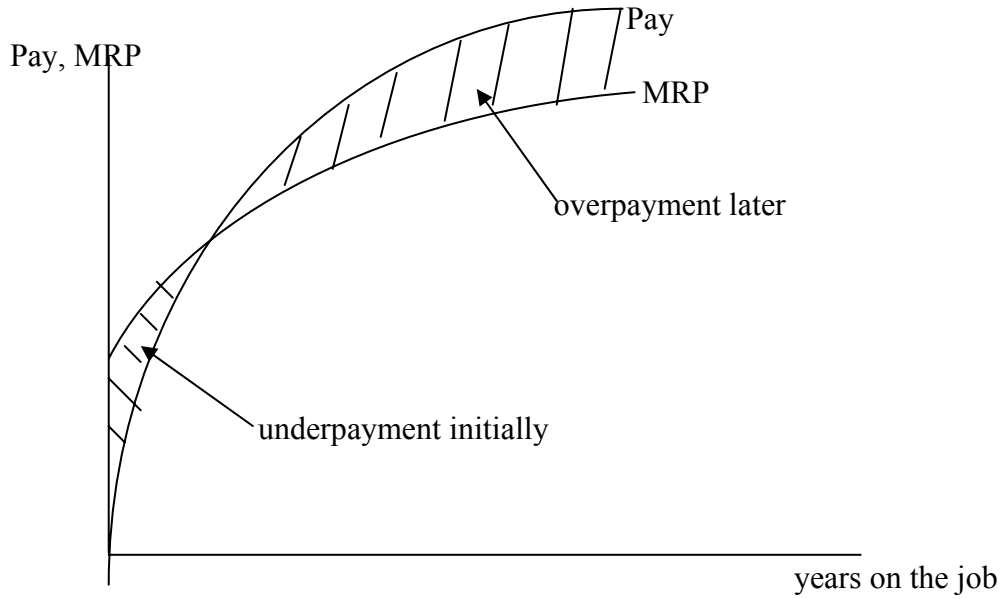
Worker shirking, whether in the form of less work effort or on-the-job theft, is enormously expensive to firms. Hence, employers have a strong incentive to offer payment schemes that encourages worker loyalty. Two such schemes are "Underpayment followed by Overpayment" (delayed compensation) which we discuss in this section, and efficiency wages which we discuss in the next lecture.

There may be reasons to think that **Underpayment followed by Overpayment** contracts may be preferred to contracts that pay the value of marginal product in each period. ES give some of these reasons on p. 401 in their text:

> 1) the Underpayment-followed-by-Overpayment contracts are attractive to those workers interested in long-term relationships and who will work diligently enough to avoid being fired---hence, in the absence of information about workers' diligence, the deferred compensation scheme can be a signaling mechanism to sort workers;

2) incentives to work diligent are increased, in order that the worker quality for the overpayments received later on, (the costs of being fired if caught shirking are greater). If such schemes are successful in getting workers to work harder over their lifetimes, then their pay will be larger than it would be for those who in firms that only pay the value of their marginal product (do you understand why would the overall, discounted present value of income may be larger?).

The diagram for the Underpayment followed by Overpayment contract would look something like the following:



Such contracts have to offer at least the same discounted present value as contracts which paid the workers their marginal revenue product (MRP) each period. That is, a worker with a discount rate of r would be indifferent between a Underpayment followed by Overpayment contract and a contract that paid their MRP each period as long as the discounted present value of the two income streams was the same (and the worker had no plans to shirk):

$$\sum_{t=0}^{T} \frac{w_t}{(1+r)^t} = \sum_{t=0}^{T} \frac{MRP_t}{(1+r)^t}$$

But if the delayed compensation scheme reduces shirking, than the overall MRP rises each period and the net present value of earnings will rise (or other firms, offering the same type of contracts, will bid away workers). In effect, a worker posts a bond with the firm during the initial years on the job, and the bond is repaid (with interest) in the later years.

Why are such schemes more likely to be offered by firms that are stable and financially secure?